Session 17 Overview:
Technologies for Human Interaction & Health
TECHNOLOGY DIRECTIONS SUBCOMMITTEE

Session Chair: ***Patrick Mercier***
University of California,
San Diego, La Jolla, CA

Associate Chair: ***Shuichi Nagai***
Panasonic, Moriguchi, Osaka, Japan

Subcommittee Chair: Makoto Nagata, Kobe University, Kobe, Japan

Innovations in sensors, actuators, computation, and communication circuits are enabling new and exciting applications in human interaction and healthcare. This session begins with an invited system paper that describes how advances in artificial intelligence and processors, networks, and actuators have enabled new classes of human-interacting robots. The next three papers then describe innovations in voice-activity detection, large-scale tactile sensing, and low-power embedded motion detection systems for human-interactive applications. Subsequent papers introduce ultra-low-power body-area networking circuits, dual optical/electrical tomographic imaging circuits, and ultra-low-power PPG sensing for health-related applications.

INVITED PAPER **1:30 PM**

17.1 AI x Robotics: Technology Challenges and Opportunities in Sensors,Actuators, and Integrated Circuits
M. Fujita, Sony, Tokyo, Japan
In Invited Paper 17.1, Sony presents key technological developments in processors, sensors, actuators, and wireless communications towards innovative new applications in AI and robotics. These are combined in a demonstration robotic platform called, aibo, which is a fully autonomous entertainment robot embedded with AI-based behavior control.

2:00 PM

17.2 A 142nW Voice and Acoustic Activity Detection Chip for mm-Scale Sensor Nodes Using Time-Interleaved Mixer-Based Frequency Scanning
M. Cho, University of Michigan, Ann Arbor, MI
In Paper 17.2, the University of Michigan demonstrates a voice and acoustic activity detector using mixer-based architecture and an ultra-low power neural network-based classifier. Due to 4× and 12× power reduction in feature extraction and neural processing, a 91.5%/90% speech/non-speech hit rate at 10dB SNR with power consumption of only 142nW is achieved.

2:30 PM

17.3 Hybrid System for Efficient LAE-CMOS Interfacing in Large-Scale Tactile-Sensing Skins via TFT-Based Compressed Sensing
L. E. Aygun, Princeton University, Princeton, NJ
In Paper 17.3, Princeton University introduces an architecture for high-spatial-resolution tactile-sensing skins based on LAE-CMOS hybrid systems with compressed sensing. By readout and control with a resistive force-sensing array using ZnO TFTs, an acquisition error of 0.7kΩ_{rms} over a 100kΩ-to-15kΩ sensing range, at energy and rate of 1.2μJ/frame and 31fps is realized.

 978-1-5386-8531-0/19/$31.00

3:15 PM

17.4 16MHz FRAM Micro-Controller with a Low-Cost Sub-1µA Embedded Piezo-Electric Strain Sensor for ULP Motion Detection

S. Khanna, Texas Instruments, Dallas, TX

In Paper 17.4, Texas Instruments presents a 0.13µm 16MHz FRAM micro-controller with an embedded piezoelectric strain sensor using an array of thin film Lead Zirconium Titanate (PZT) capacitors for motion detection. The sensor requires only 650nA and occupies an area of 0.6mm^2.

3:30 PM

17.5 A 0.8mm^3 Ultrasonic Implantable Wireless Neural Recording System with Linear AM Backscattering

M. M. Ghanbari, University of California, Berkeley, CA

In Paper 17.5, the University of California, Berkeley describes a 0.8mm^3 neural recording system that is powered by and communicates via ultrasound. A linear AM backscatter technique is introduced that enables operation at a depth of up to 5cm while operating a small power-efficient LNA with –44dB THD over a 20mV$_{pp}$ input range.

3:45 PM

17.6 A Sub-40µW 5Mb/s Magnetic Human Body Communication Transceiver Demonstrating Trans-Body Delivery of High-Fidelity Audio to a Wearable In-Ear Headphone

J. Park, University of California, San Diego, La Jolla, CA

In Paper 17.6, the University of California, San Diego presents a transceiver design to exploit the low path loss enabled by magnetic human body communication (mHBC). A test chip demonstrates reliable trans-body communication at 5Mb/s at TX and RX power consumptions of 18.56 and 6.3µW, respectively.

4:15 PM

17.7 A 7.0fps Optical and Electrical Dual Tomographic Imaging SoC for Skin-Disease Diagnosis System

Y. Lee, KAIST, Daejeon, Korea

In Paper 17.7, KAIST describes a portable lightweight (3g) system that can simultaneously image the skin in the electrical and optical domains to help improve skin disease diagnostics. Incorporation of diffused optical and electrical impedance tomography enables achieving 0.5mm sensitivity and 2mm selectivity.

4:45 PM

17.8 A 2.6µW Monolithic CMOS Photoplethysmographic Sensor Operating with 2µW LED Power

A. Caizzone, EPFL, Neuchâtel, Switzerland

In Paper 17.8, EPFL presents a photophlethysmography (PPG) sensor that utilizes an array of dedicated pinned-photodiodes in a 0.18µm CIS process. Thanks to improved sensitivity offered by the diodes, the total power of the PPG sensor is 4.6µW.

17.1 AI x Robotics: Technology Challenges and Opportunities in Sensors, Actuators, and Integrated Circuits

Masahiro Fujita

Sony, Tokyo, Japan

In 1956 at the Dartmouth conference, the terminology of artificial intelligence (AI) was first used. In those days it was also called as symbolic AI (Symbolic-AI) [1]. For example, let us assume a block world problem in Fig 17.1.1 (right), where a block is represented as a symbol, "Block1", and it can be operated by operators such as "PICKUP(Block1). In order to apply the operator "PICKUP" to the target object, "Block1", the AI system has to check the pre-condition such as "CLEAR Block1", which means there is no object on "Block1". The Fig.17.1.1 (right) shows an example of a task from State-A to State-B. The system has to search the possible operators and the pre-conditions so that State-B is achieved. There are many basic algorithms developed in Symbolic-AI era, which are often used today including the A*-search algorithm. Shakey is the representative example of intelligent robots based on Symbolic-AI. It was a wheel-based movable robot equipped with a TV-camera, Laser-Range-Finder, etc. It can move blocks in the real world using Symbolic-AI technologies. Its behavior control architecture is shown in Fig 17.1.1 (left). It has three steps, SENSE, PLAN, and ACT. Therefore, it is known as the SENSE-PLAN-ACT architecture. It is computationally intensive especially in the PLAN computation, therefore it is difficult if the environment is dynamically changing.

To solve the issue of Symbolic-AI, in the 1980s, Embodied-AI, or Behavior-Based-AI, was developed [2]. Its behavior control architecture is shown in Fig 17.1.2 (left). Compared with Symbolic-AI, it is composed of multiple layers, and basically there is no PLAN step. Genghis was the representative of Embodied-AI, which looked like a 6-legged insect, and had two feelers [2], It didn't symbolize the environment, nor search in the symbolic world, but sensed the environment and reacted according to sensing information. Thus, it was also called SENSE-ACT architecture compared with SENSE-PLAN-ACT architecture. For example, in the Avoiding Action layer, if the left feeler becomes on, then turn-right action is activated so that it can avoid the obstacles on the left. In the Wondering Action layer, if there are no stimuli of the feelers, then the go-forward action is activated and after a while the random-turn action is activated. In the Exploring Action layer, using the distance sensor such as an ultra-sonic sensor, it activates the wall-following action so that it explores the world by move along walls, which is considered a good strategy for exploration in a maze. Avoiding Action's priority is higher than Wondering Actions, and Wondering Action's priority is higher than Exploring Action. Thus, according to the complex environment, Genghis shows complex behaviors without planning. However, it is difficult to achieve intelligent behaviors using such as planning.

In the 1990s, Hybrid-AI was developed that integrated both of Symbolic-AI and Embodied-AI [2]. Hybrid-AI's behavior control architecture is shown in Fig 17.1.2 (right). There are three pathways in the behavior control architecture. One is the symbol representation pathway to the deliberative behavior layer, the second is the situated base pathway to situated behavior layer, and the third is a quick response pathway to reflexive behavior layer. In the deliberative layer, the Symbolic-AI such as planner is executed, and in the situated behavior layer, many behavior modules are evaluated and coordinated to execute proper behavior according to the situation. In the reflexive behavior layer the action is directly activated as a reaction of external stimuli in an emergent situation such as avoiding a collision. In the situated behavior layer the coordination of behaviors is often a selection of the most proper one, or mixed if the robot can do those in parallel. The behaviors from deliberative behavior layer are usually to give bias signals to the behaviors in the situated behavior layers.

The behavior architecture of an entertainment robot AIBO released in 1999 was based on Hybrid-AI, but in addition it has internal states so-called pseudo instinct [4]. AIBO's behavior control architecture is shown in Fig 17.1.3. In AIBO it evaluates the behaviors based on both the external states and the internal states. The dynamics of the internal states is 1) decreasing as time proceed, 2) increasing when related behavior to the internal state is executed successfully. The evaluation score of the behavior is high when the related internal state is out of a set range. For example, if the internal state named exercise-instinct is bellow the low limitation of the range, the evaluation score of behaviors that can increase the exercise-instinct such as ball-kicking behavior becomes high. In this case, if a ball is found near AIBO, which means the external states also support the evaluation score of ball-kicking behavior. Then, if ball-kicking behavior is executed to increase the exercise-instinct and to be in the range. Thus, it is a homeostasis system and it is a base mechanism that AIBO can execute spontaneous behaviors without any commands from users.

Let us compare AIBO with a new aibo released in 2018. Hereafter we distinguish the two robots using AIBO released in 1999 and aibo released in 2018. The hardware architectures of AIBO and aibo are shown in Fig 17.1.4 (left) and Fig 17.1.4 (right). Regarding the processor, AIBO used a RISC architecture, a MIPS R400x with about 300MHz clock. There is a companion chip with functions, 1) image filter bank for pyramidal image generator, 2) object detection based on 8 colors, 3) Inner product engine, and 4) Serial bus manager for devices. The combination of the R400x and the companion chip makes it possible to process images and other sensors in real time. On the other hand, aibo uses a 2.4GHz Snapdragon SoC found in smartphones. The SoC has 4 ARM CPU cores, DSPs, GPUs, etc, and using only the SoC it can process multiple imagers and other sensors in real time. Regarding the wireless LAN (WLAN), AIBO uses 802.11b (22.2Mbps max), but aibo has 802.11n (77.2Mpbs max), and cell-phone network LTE(150Mbps max). One of the big advantages of aibo is in the progress of the wireless communication environment enabling software updates that can be done through the cloud. In the AIBO case, it should be done using a physical memory card. In addition, the power of cloud computing and storage is very useful for the embedded system to off-load computation tasks such as learning, to keep important state information of aibo, etc. In addition, the wide availability of smartphone helps the user to control and monitor the aibo using a graphical user interface (GUI).

More details of sensors, actuators, effectors, and AI processing of aibo are shown in Fig 17.1.5. and Fig 17.1.6. Two back-illuminated(BI) -CMOS sensors are used in aibo. One CMOS sensor is at the top of the nose, and its image signal is used to detect human faces, hands, legs, and torsos using an embedded deep neural network. In addition, aibo detects the objects that designed for aibo to play with or use, such as the pink ball, the pink born like object named ai-born, the charge station, AIBO, and aibo. Furthermore, aibo can detect general objects such as table, chair, TV, which are usually found in a living room. Another CMOS sensor is in its back. It is used for so-called simultaneous localization and mapping (SLAM), that is to build a map of a room by moving and to infer a location and posture of aibo itself.In the front face and breast, there equipped a time-of-flight (ToF) sensor for obstacle detection, pyroelectric-IR sensor for live human detection (not human in photo), and position sensing detection (PSD) sensor for near-located-obstacle and a cliff such as table edge. There are 4 microphones to estimate the sound source direction, and to reduce noise using beamforming for improving speech recognition performance. Speech recognition is processed in the SoC using a deep neural network for acoustic modeling. There are three tactile sensors in the head, chin, and back to engage the physical interaction with users such as patting for reward. Two inertia meter unit (IMU) are equipped at the head and the body for utilizing with visual SLAM, and for detecting when held up by users. There are 22 actuators to realize real dog-like motions including mouth and ears. Those actuators are controlled by a servo-control MCU in the body. Its eyes are displayed using two OLED devices. It is important for aibo to show emotions using the eye patterns such as happy, angry, stressed, sleepy, and surprise, etc.

As CPUs, networks, and wireless communications progress, so will AI and Robotics progress. Three examples are mentioned here. First one is about imager. Imagers were originally developed for taking a scene like photos and movies that are watched by humans. Now, imagers are developed for other applications such as perception and recognition by machines. For example, a BI stacked CMOS image sensor can take a video with 1000 frame/sec and can track objects with logic circuits on the back-side[6]. It is more capable in ability than human perception. This direction will be more enhanced with cloud computing. Second one is about force control. In order to realize a human-robot collaborative society, safety physical interaction is necessary. Force sensors and tactile sensors with force control is important [7]. Thirdly, multiple robot systems integrated with a cloud system are important in the creation of a human-robot collaborative society. Finally, wireless communication with wide bandwidth, low latency, and secure protocol is an important technology.

References:

[1] A.Russel and P.Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 1995.

[3] R. Murphy and R. Arkin, *Introduction to AI Robotics*, MIT Press, 2000.

[4] M. Fujita, et al., "An open architecture for robot entertainment", *Autonomous Agents*, pp. 435-442, 1997.

[5] M. Fujita, "Intelligence Dynamics: a concept and preliminary experiments for open-ended learning agents", *Journal of AAMAS*, vol. 19, no. 3, pp. 248-271, Dec. 2009.

[6] T. Yamazaki, et al., " A 1ms High-Speed Vision Chip with 3D-Stacked 140GOPS Column-Parallel PEs for Spatio-Temporal Image Processing" *ISSCC Dig. Tech Papers,* pp. 82-83, Feb. 2017.

[7] K.Nagasaka, et al.,"Whole-body Cooperative Force Control for a Two-Armed and Two-Wheeled Mobile Robot Using Generalized Inverse Dynamics and Idealized Joint Units", ICRA2010, pp. 3377-3383, 2010.

 978-1-5386-8531-0/19/$31.00

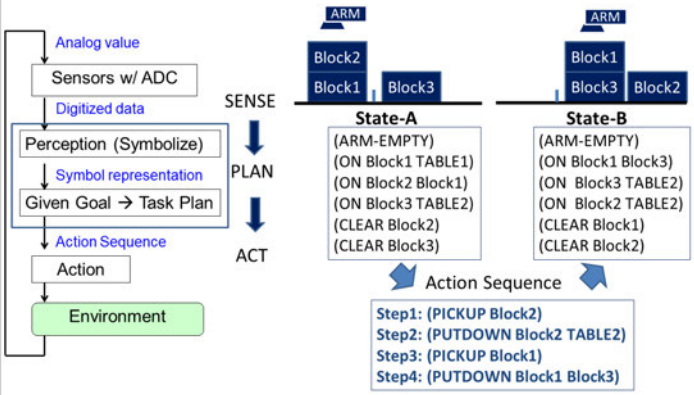


Figure 17.1.1: Symbolic-AI behavior control architecture (left) and typical representation of block world (right). Symbolic-AI is often called Sense-Plan-Act, and the Plan part is computationally intensive.

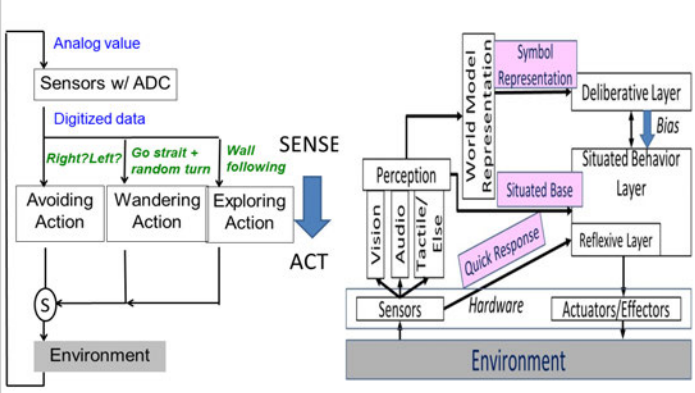


Figure 17.1.2: Embodied-AI behavior control architecture (left) and Hybrid-AI behavior control architecture (right). The notation of "S in a circle" means that actions from multiple layers are properly selected.

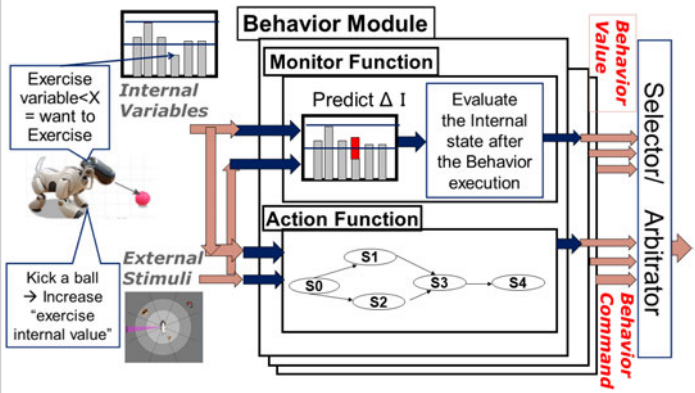


Figure 17.1.3: AIBO behavior control architecture based on Hybrid-AI (left) and action evaluation of AIBO using both external stimuli and internal stimuli (pseudo instinct) (right).

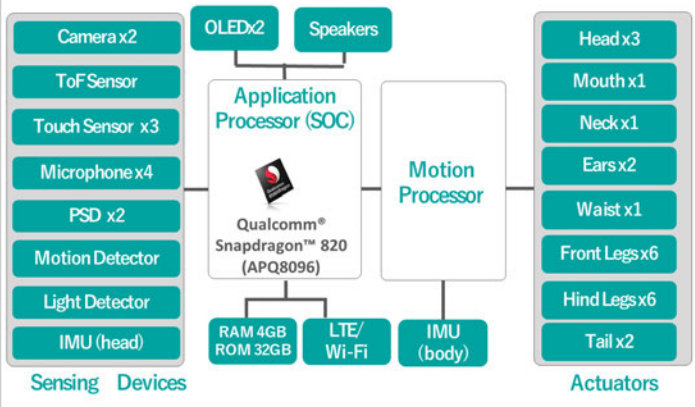


Figure 17.1.4: Hardware architecture of aibo. AIBO needed a dedicated LSI for visual processing, Inner Product Engine, etc., and multiple servo-controller for actuators. In the other hand, aibo needs one SoC and Motion Controller for all actuators.

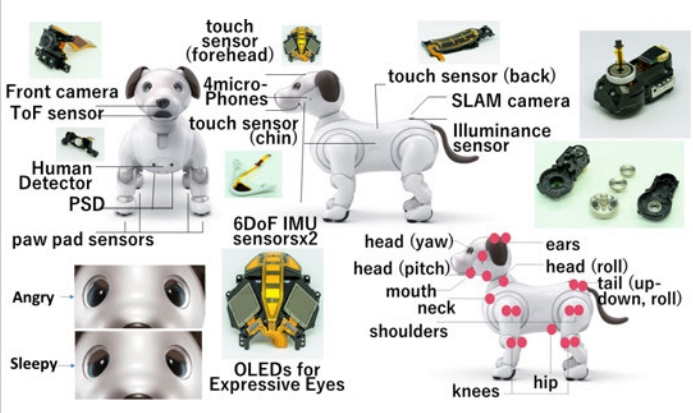


Figure 17.1.5: Major sensors and effectors of aibo.

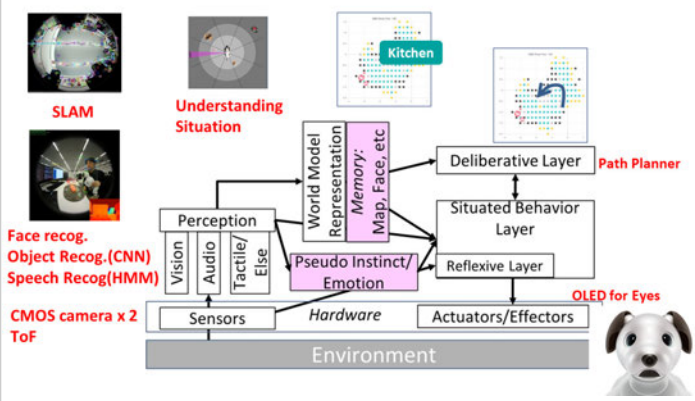


Figure 17.1.6: aibo's behavior control architecture. Left bottom picture is the output of the deep neural network for face and object detection. The top pictures from left to right are, camera output in back for SLAM, memory of map in the room with label "Kitchen" and path planner result from the current position to "Kitchen".

17

17.2 A 142nW Voice and Acoustic Activity Detection Chip for mm-Scale Sensor Nodes Using Time-Interleaved Mixer-Based Frequency Scanning

Minchang Cho*, Sechang Oh*, Zhan Shi, Jongyup Lim, Yejoong Kim, Seokhyeon Jeong, Yu Chen, David Blaauw, Hun-Seok Kim, Dennis Sylvester

University of Michigan, Ann Arbor, MI
*Equally-Credited Authors (ECAs)

Acoustic sensing is one of the most widely used sensing modalities to intelligently assess the environment. In particular, ultra-low power (ULP) always-on voice activity detection (VAD) is gaining attention as an enabling technology for IoT platforms. In many practical applications, acoustic events-of-interest occur infrequently. Therefore, the system power consumption is typically dominated by the always-on acoustic wakeup detector, while the remainder of the system is power-gated the vast majority of the time. A previous acoustic wakeup detector [1] consumed just 12nW but could not process voice signals (up to 4kHz bandwidth) or handle non-stationary events, which are essential qualities for a VAD. Prior VAD ICs [2,3] demonstrated reliable performance but consumed significant power (>20μW) and lacked an analog frontend (AFE), which further increases power. Recent analog-domain feature extraction-based VADs [4,5] also reported μW-level power consumption, and their simple decision tree [4] or fixed neural network-based approach [5] limited broader use for various acoustic event targets. In summary, no sub-μW VAD has been reported to date, preventing the use of VADs in unobtrusive mm-scale sensor nodes.

This work presents a 142nW programmable, neural network-based acoustic sensing system for both VAD and non-voice event detection. We use a *time-interleaved mixer-based architecture* that sequentially scans and down-converts the 4kHz bandwidth signal to a ≤500 Hz passband, reducing amplifier, ADC, and DSP power by 4×. The neural network (NN) processor employs computational sprinting, which minimizes static energy dominance in low frequency/voltage regime, providing 12× power reduction in the digital domain. The architecture (Fig. 17.2.1, top) has two signal chains: an ULP chain with 142nW consumption that is always on and a 18μW high performance (HP) chain that wakes upon event detection by the ULP chain. Unlike the ULP chain, the HP chain has a full 4kHz bandwidth AFE while sharing the same digital backend with the ULP chain. In addition to VAD, the system features an inaudible acoustic signature detection mode to enable remote silent system wakeup. With always-on VAD, the system has a 4.5-year lifetime with a 5mm mini coin-cell battery (2mAh) and achieves 91.5% voice detection accuracy.

Figure 17.2.1 shows the *time-interleaved mixer-based architecture* that reduces power consumption of AFE and DSP by lowering their bandwidth and sampling rate to 500Hz and 1kHz, respectively. The incoming signal from the microphone is amplified by an LNA with the full 4kHz bandwidth. At this point the mixer, switched by a binary discrete cosine transform (DCT) sequence, immediately down-converts the frequency of a desired feature to a programmable intermediate frequency (IF) of <500Hz. The digital binary sequence generator supports an arbitrary DCT frequency for the mixer switch control; for example, the 4kHz band can be divided into 31.25Hz frequency bins using a 128-pt DCT, and the energy content of 32 bands is sequentially extracted by sweeping DCT frequencies (F_1, … F_{32}). The 32 bands are chosen during NN training for each target event. The IF down-converted signal is further amplified and low-pass filtered with 500Hz bandwidth (via a PGA) and digitized at 1kS/s. Finally, the digital IF mixer down-converts the signal to DC, and the feature power is measured. With a DCT length of 16ms per feature (128-pt DCT with 8kHz binary mixing), 32-feature extraction requires a 512ms frame. The mixer-based structure reduces bandwidth, sampling rate, and clock frequency of AFE and DSP after the mixer; thus, the feature extraction power consumption is decreased from 225nW (simulation) to 60nW. IF is set to ~250Hz to avoid PGA 1/f noise, while the image aliasing issue of non-quadrature mixing is mitigated by a NN trained with image-aliased signals.

Figure 17.2.2 shows the circuit diagram of the AFE with ULP and HP chains. Each chain consists of an LNA, PGA, ADC driver, and ADC. Both chains share a single MEMS microphone and charge pump. For high sensitivity, a three-stage Dickson pump is used to bias the microphone at 10V. The microphone switches between the chains by controlling *ULP_CH_EN* and *HP_CH_EN*, which are level-shifted to 10V. As coupling capacitors for the level-shifting may suffer from leakage between infrequent mode switches, we refresh their charge periodically. Capacitive feedback and pseudo-resistor dc-servo loops are used for the ULP LNA (18dB gain). The LNA OTA adopts an inverter-based cascode amplifier for noise efficiency. Its common-mode feedback is composed of two loops. Coupling capacitors provides fast loop response, and the DDA output sets the DC voltage. The auxiliary amplifiers (Aux-amp) in the dc-servo loop shift common-mode voltages for high dynamic range. In addition, Aux-amp attenuates large LNA output due to diode-connected nature, which reduces the maximum amplitude seen by the pseudo-resistors, reducing their amplitude-dependent drift. The mixer is composed of transmission gates switched by the DCT sequence generator. Unlike the LNA OTA, PGA OTA uses only a PMOS input pair for the maximum output range. By tuning cap C_{I2}, the gain is adjustable between 4.5 and 31.2dB, and C_L sets 500Hz BW for ULP mode. The ADC driver is followed by an 8b SAR ADC. The HP blocks are similar to the ULP counterparts except that they are scaled for low noise and full 4kHz bandwidth. We minimize the ULP-HP transition time by temporarily turning on the fast settling switches during the transition. This helps to set the common-mode voltage very quickly (100ms vs 6s, measured).

Figure 17.2.3 (top) depicts the digital backend architecture. Always-on modules are implemented with thick oxide I/O devices to suppress leakage, while power-gated modules (NN processor, FIFO, and audio compressor [6]) are designed with standard devices. The NN processor uses a 16kB custom ultra-low retention leakage SRAM for the 4-bit weights storage, and its ISA includes matrix-vector multiplication, non-linear activation, FFT, conditional branch, element-wise vector operation, and min/max/averaging. In the ULP mode, the processor sprints at a relatively high frequency clock (700kHz, Fig. 17.2.3, bot right) when the sequential feature extraction is complete (every 512ms) and then power-gates to minimize the leakage power, resulting in 12× power reduction (sprint/sleep ratio of 0.008). In HP mode, the processor computes a full (non-sequential) FFT and a larger NN without duty cycling for the improved latency (32ms) or hit rate at the cost of higher power consumption (14μW). The binary mixer sequence generator (Fig. 17.2.3, bot left) is programmable for different DCT sizes, feature frequency resolutions, and number of features. Due to the mixer-based architecture, digital processing runs at 1kHz (vs. 8kHz Nyquist rate), yielding 41% reduction of feature extraction power.

The system also features inaudible acoustic signature detection to enable silent remote system wakeup (Fig. 17.2.4). The binary mixing sequence is replaced with a maximal length sequence (MLS) signature generated by a 1kHz programmable LFSR. Correlation between the incoming wakeup signal and the local sequence is performed through the ULP mixer and PGA. To synchronize the wakeup and local sequence, we employ a time-drift synchronization scheme that uses intentional frequency mismatch between the two so that they naturally time-synchronize periodically. This inaudible (–10dB SNR) signature detection consumes only 66nW with 4s worst case latency.

The chip is fabricated in 180nm CMOS and integrated with a MEMS microphone (Fig. 17.2.7). The ULP and HP chain amplifiers consume 31nW and 370nW with 16μVrms and 9.1μVrms input-referred noise, respectively. Figure 17.2.5 (top left) shows the measured mixer-based frequency scanning operation and input referred noise spectrum. Figure 17.2.5 (bot left) shows the measured ULP chain power breakdown. For VAD evaluation, speech from the LibriSpeech dataset is mixed with babble noise from the NOISEX-92 dataset. NN training and evaluation use exclusive datasets. Figure 17.2.6 compares the system with prior work. The system achieves 91.5%/90% speech/non-speech hit rates at 10dB SNR with babble noise (electrical test, Fig. 17.2.5 top right) in ULP mode when programmed with a NN of size 32-32-16-2 neurons, exhibiting ~7.5% better hit rate at 7× less power consumption than prior state-of-the-art. Unlike prior-art, we also report acoustic VAD test results measured in a sound chamber, showing >83%/85% speech/non-speech hit rates with a signal level down to 50dBA SPL (Fig. 17.2.5, bot right).

Acknowledgements:
This work was supported by Defense Advanced Research Projects Agency (DARPA). The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References:
[1] S. Jeong, et al., "A 12nW Always-on Acoustic Sensing and Object Recognition Microsystem Using Frequency-Domain Feature Extraction and SVM Classification," *ISSCC Dig. Tech. Papers*, pp. 362-363, Feb. 2017.
[2] A. Raychowdhury, et al., "A 2.3 nJ/Frame Voice Activity Detector-Based Audio Front-End for Context-Aware System-On-Chip Applications in 32nm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 8, pp. 1963-1969, Aug. 2013.
[3] M. Price, et al., "A Scalable Speech Recognizer with Deep-Neural-Network Acoustic Models and Voice-Activated Power Gating," *ISSCC Dig. Tech. Papers*, pp. 244-245, Feb. 2017.
[4] K. Badami, et al., "Context-Aware Hierarchical Information-Sensing in a 6μW 90nm CMOS Voice Activity Detector," *ISSCC Dig. Tech. Papers*, pp. 430-431, Feb. 2015.
[5] M. Yang, et al., "A 1μW Voice Activity Detector Using Analog Feature Extraction and Digital Deep Neural Network," *ISSCC Dig. Tech. Papers*, pp. 346-347, Feb. 2018.
[6] M. Cho, et al., "A 6×5×4mm^3 General Purpose Audio Sensor Node with a 4.7μW Audio Processing IC," *IEEE Symp. VLSI Circuits*, pp. 312-313, June 2017.

 978-1-5386-8531-0/19/$31.00

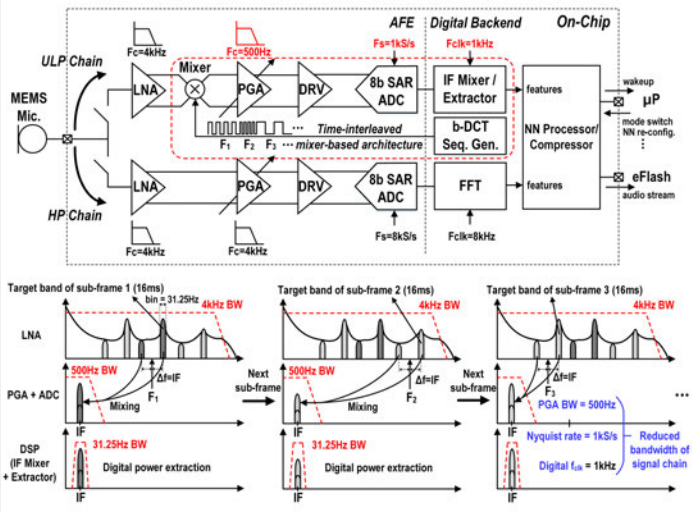


Figure 17.2.1: Acoustic sensing system architecture (top), and operation principle of time-interleaved mixer-based frequency scanning (bottom).

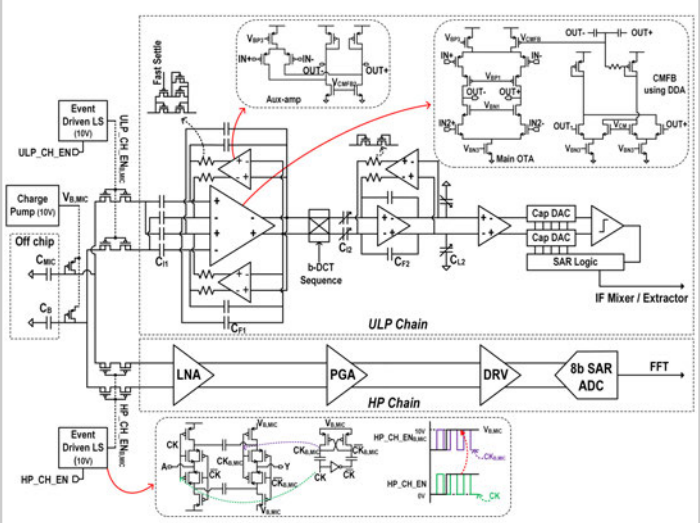


Figure 17.2.2: Circuit diagram of the analog front-end with ULP and HP chains.

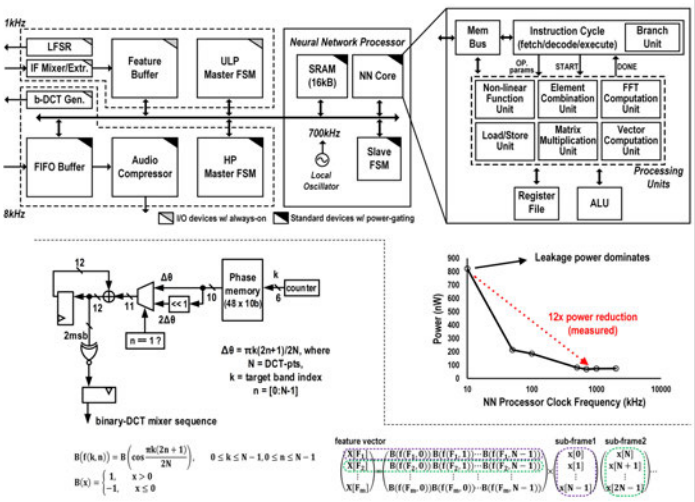


Figure 17.2.3: Digital backend architecture including neural network processor (top), measured power reduction from computational sprinting (bottorm right), and binary DCT mixer sequence generator (bottom left).

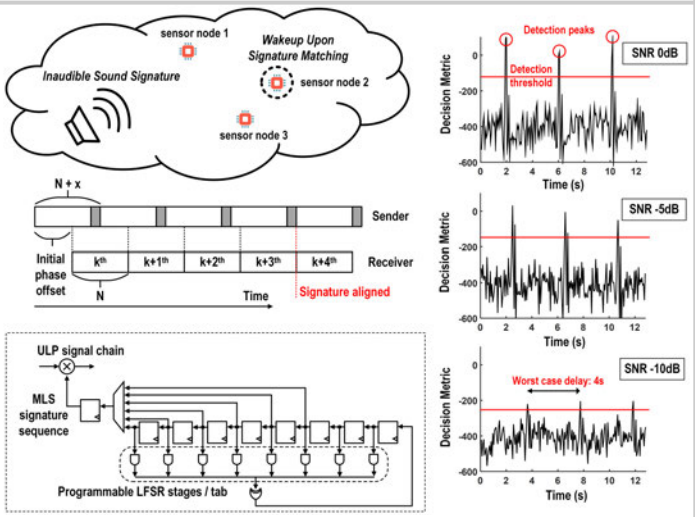


Figure 17.2.4: Acoustic signature wakeup detection (left), and measurement results with 6 stages, 63-length sequence at various SNRs (right), showing detection down to -10dB SNR.

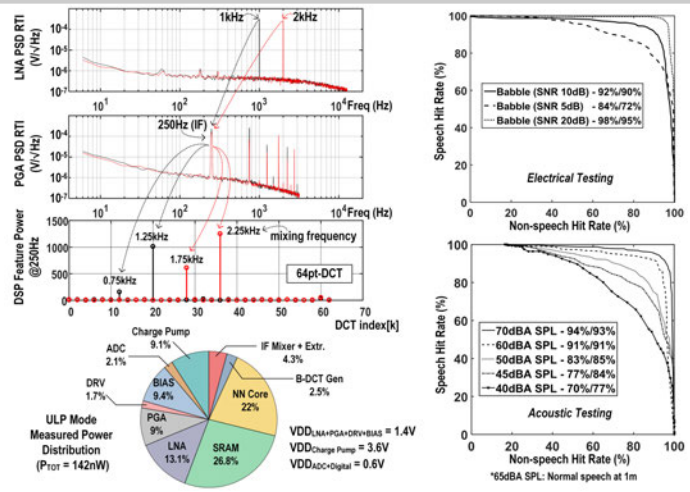


Figure 17.2.5: Chip measurement results. Power spectral density for LNA, PGA, and DSP (top left). Two different applied tones are mixed down to 250Hz in IF and extracted by DSP at two mixing frequencies each. ULP mode power distribution (bot left), and ROC curves for VAD (right).

Feature Extractor	This Work	ISSCC'18 [5]	ISSCC'16 [1]	ISSCC'15 [4]
Technology (nm)	180	180	180	90
Feature Extraction Type	Mixed-signal	Analog to events	Digital	Analog
Channel Number	16 – 48	16	4 – 16	16
Frequency Range (Hz)	75 – 4k	100 – 5k	0.2 – 470	75 – 5k
Power (nW)	60	380	10	6000
Normalized Power* (nW)	5	71	34	1186
Dynamic Range (dB)	47	40	N/A	40
Building Blocks	LNA, Mixer, LPF, DSP	LNA, BPF, FWR, IAF	DSP	LNA, BPF, FWR, LPF

ULP AFE(LNA+PGA) Summary	
Gain(dB)	22.5-49.2
Out Noise (mVrms)	0.83@min gain 4.6@max gain
Max Output (Vpp)	1.45(<0.4%THD)
Dynamic Range(dB)	55.8@min gain 40.9@max gain
Power (nW)	31
Bandwidth (Hz)	500
VDD(V)	1.4

Voice Activity Detector	This Work	ISSCC'18 [5]	ISSCC'17 [3]	ISSCC'15 [4]	JSSC'13 [2]
Technology (nm)	180	180	65	90	32
Acoustic Input	Analog/passive mic. w/ gain stage	Analog mic. w/ gain stage	Assume digitized	Analog mic. w/ gain stage	Assume digitized
Classifier	Neural network	Neural network	Neural network	Decision tree	Energy-based
Classifier Topology Programmability	Yes	No	Yes	No	No
Dataset**	LibriSpeech + NOISEX-92	AURORA4 + DEMAND	AURORA2	NOISEUS	N/A
Latency (ms)	512	10	10	<100	10
Power (μW)	0.142	1	22.3	6	~300
Accuracy SP/Non-SP hit rate (electrical test)	91.5%/90%*** @ babble 10dB SNR	84%/85% @ restaurant 10dB SNR	90%/90%**** @ unspecified context 7dB SNR	89%/85% @ babble 10dB SNR	97% accuracy @ unspecified context unspecified SNR
Acoustic Testing Performed	Yes	No	No	No	No

*calculated according to the equation in [5], normalized to 4kHz; **all datasets are similar in speech quality;
***measured at ULP mode with 128pt-DCT, 32 feature channels, and 250Hz IF. ****converted from EER in [3].

Figure 17.2.6: Comparison table for feature extractor (top left), VAD (bottom), and performance summary of ULP AFE (top right).

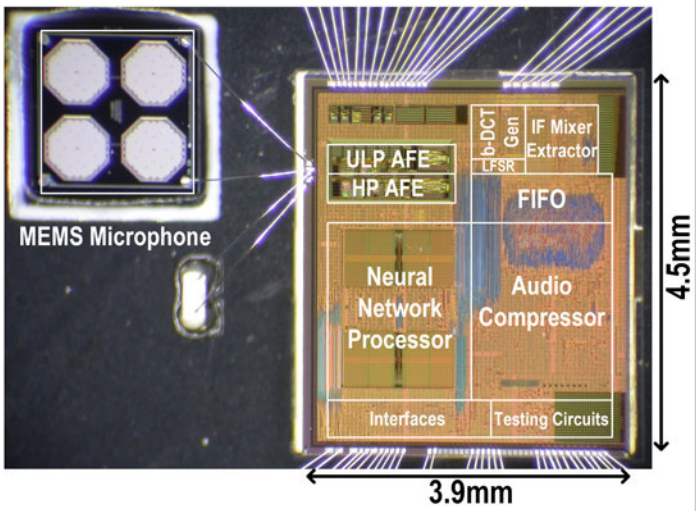


Figure 17.2.7: Die micrograph and system intgration with MEMS microphone.

978-1-5386-8531-0/19/$31.00

17.3 Hybrid System for Efficient LAE-CMOS Interfacing in Large-Scale Tactile-Sensing Skins via TFT-Based Compressed Sensing

Levent E. Aygun*, Prakhar Kumar*, Zhiwu Zheng*, Ting-Sheng Chen*, Sigurd Wagner, James C. Sturm, Naveen Verma

Princeton University, Princeton, NJ
*Equally-Credited Authors (ECAs)

Tactile sensing has wide-ranging applications, from intelligent surfaces to advanced robotics. Large-Area Electronics (LAE), based on low-temp. fabrication (<200°C) of thin films, presents distinct capabilities, due to compatibility with a broad range of materials (enabling diverse transducers), as well as large and flexible substrates and materials-deposition methods (enabling expansive and formfitting sensing arrays). However, low performance/energy-efficiency of LAE thin-film transistors (TFTs) necessitates hybrid systems, integrating Si-CMOS ICs for system functions (sensor readout/control, processing, etc.). Initial work shows that a primary challenge in hybrid systems is the large number of interfaces required between LAE and CMOS, particularly as the number of sensors scales [1,2]. This paper presents a force-sensing system that exploits signal sparsity exhibited in many large-area tactile-sensing applications (e.g., detecting point damage/stress in structures [3]), to reduce interfacing complexity to the level of sparsity, rather than a level related to the number of sensors (e.g., [1]). This is achieved via compressed sensing (CS), enabling sensor-acquisition by simple switches, readily implemented using TFTs. While CS has previously been leveraged in a hybrid-system architecture targeting signal sampling-rate requirements [2], this system applies it for high spatial resolution in tactile sensing.

Figure 17.3.1 summarizes CS and the system-level motivations. As shown, CS involves sampling different superpositions $\vec{y}$ of input data $\vec{x}$ (having dimensionalities M and N, respectively), via a measurement matrix ϕ. CS theory says that: (1) if $\vec{x}$ has at most K non-zero elements (K sparse), only $M>K\times\log_2(N/K)$ different superpositions are needed to reconstruct $\vec{x}$; and (2) a random and simple matrix ϕ, with elements 0/1, can be found for acquiring such superpositions. This makes TFT implementation possible, and Fig. 17.3.1 compares how system complexity (vs. number of sensors N) would scale compared to traditional TFT architectures for hybrid-system interfacing (sequential scanning, active matrix), based on the following key metrics: (1) number of acquisition cycles (operational complexity); (2) number of interfaces (system-assembly complexity); (3) dynamic range (DR) of each interface (CMOS/readout complexity); (4) number of TFTs (LAE complexity). If M,K<<N (i.e., high sparsity), CS substantially benefits cycles and interfaces, which together with DR represent total data bandwidth of transfer, and can be traded-off. But, modest increase in DR (~K) is preferred, as it pushes complexity to the high-efficiency CMOS domain and because sensor proximity to embedded signals in tactile sensing typically yields large responses [3]. A drawback with CS is the number of TFTs; but, as described below, the statistical nature of acquisition with CS enables high tolerance to typical TFT faults.

Figure 17.3.2 shows the CS hybrid-system architecture, demonstrated for an array of resistive force sensors (applicable to other resistive sensors, e.g., temp./strain/pressure/light). While the transfer function of $R_{SNS,i}$, shown for 10 sensors, exhibits variation, the aim of tactile sensing is often coarse readout from each sensor, but with high spatial resolution. Each sensor feeds an access TFT, controlled to implement the 0/1 elements of ϕ for one column. A TFT Matrix-Logic Control block controls the access TFT, via a Matrix-Row-Selection code R/Rb[4:0] from CMOS. The control-block details are shown, consisting of matrix TFTs arranged in 32 branches with 5 TFTs/branch. The Matrix-Control signal C/Cb[4:0] enables one branch at a time, driving the access-TFT gate high/low (V_{ON}/V_{OFF}). To simplify design, the Matrix-Logic Control block is the same for each sensor, but the connections between R/Rb[4:0] and C/Cb[4:0] are different, so that the access TFTs are enabled differently as R/Rb[4:0] is cycled through 32 codes (e.g., truth-table shown for 20th sensor). By superimposing the currents from accessed sensors via the CMOS TIA, which ensures constant sensor biasing of 0.4V, a 32-row ϕ matrix is realized. The architecture thus employs a single interface and M=32 cycles, one for each conceptual row of ϕ (same effective data bandwidth as Fig. 17.3.1); but, an additional 5 differential interfaces are required for R/Rb[4:0]. M=32 supports up to N=120 sensors at target sparsity of K=3, but 20 sensors are implemented for demonstration (requiring 20×(5×32)=3200 matrix TFTs). The W/L's for matrix/access TFTs are shown, designed for access-TFT on-resistance R_{ACC}≈1.5kΩ, well below the target range for force sensors.

Figure 17.3.3 analyzes TFT fault tolerance. While all interfacing architectures in Fig. 17.3.1 require an access TFT, the 160 matrix TFTs per sensor pose notable overhead (though, this increases slowly with number of sensors N, due to $\log_2$-dependence of M on N). Figure 17.3.3 shows substantial system tolerance to matrix TFT faults. Typical faults include: (1) S-D open; (2) S-D leakage/short; (3) G-S/D leakage/short. Probability of G-S/D shorts must be controlled, as it impacts the global control signal R/Rb[4:0]; this can be done via gate-dielectric processing (oxide quality/thickness, minimizing gate-electrode hillocks). Other faults can be analyzed by considering their effective impact on ϕ. As shown, S-D opens cause element values to be set by charge dynamically held at the access-TFT gate from the previous state, corresponding to previous row. On the other hand, S-D leakage/shorts and G-S/D leakage cause contention at the access-TFT gate, making elements take intermediate values. Monte Carlo simulations, applying such fault models, show high reconstruction SNR (RSNR) is maintained to high fault rates (much higher than typically allowed for access TFTs), and RSNR improves with even higher N. This is due to use of statistical-optimization methods and many superpositions in the CS reconstruction process.

Figure 17.3.4 shows a block diagram of the CMOS IC, having 8 channels of a TIA, offset-correction, and 10b ADC readout chain, as well as digital control and R/Rb[4:0] generation. Offset correction consists of a 32-word register file and 7b current DAC (I-DAC). This enables an offset-correction current to be applied to the TIA for each matrix-row readout, via a code in the register file determined at start-up (with no force applied). The I-DAC consists of 6 binary-weighted N/PMOS current sources. The TIA employs a 2-stage op-amp, designed for stability with 500pF of input capacitance, supporting >200 access-TFT load. The ADC is a 10b SAR with 5b main/sub cap-DACs. The 8 channels enable further scaling in sensor number, via parallel arrays.

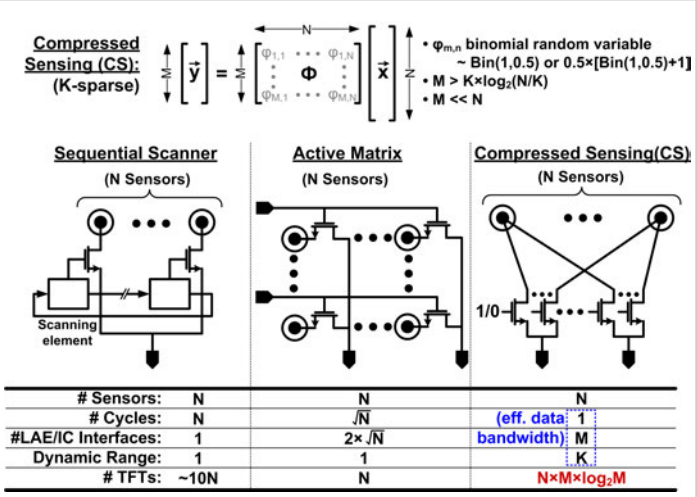
The force-sensing system is implemented with in-house-fabricated ZnO TFTs, commercial force sensors, and custom 130nm CMOS IC (Fig. 17.3.7). To ease testing, 3 PCBs are used, consisting of: (1) 20 force sensors; (2) 20 wire-bonded matrix-control/access-TFT die; (3) CMOS IC. The TFTs are fabricated on glass for dicing and wire bonding, but are fully flex compatible (process temp. <200°C). Figure 17.3.5 (left) shows overlaid waveforms from 32 TFT die from one sample, showing proper matrix-TFT operation, and proper pull up/down by access TFT (with test 10kΩ load resistor connected to 1V). Also shown are the I_D-V_{DS}'s of the access TFTs, showing $1k\Omega<R_{ACC,i}<1.9k\Omega$ (in V_{DS} range of interest for sensor resistance). Figure 17.3.5 (right) shows CMOS IC measurements, including ADC and full-system (TIA+ADC) DNL/INL, as well as the ADC code vs. I-DAC code (no input current). A summary table is at the bottom.

Figure 17.3.6 shows the demonstration setup, as well as full-system sensor acquisition/reconstruction. A sample heat map is shown for force-sensor resistance, both directly measured and reconstructed by the system, with <2.9% error (after correcting for nominal $R_{ACC,i}$=1.5kΩ). The scatter plot is derived from many such measurements, showing error of $0.7k\Omega_{RMS}$. R/Rb[4:0] codes switch at 1kHz, giving frame rate of 1kHz/32=31fps, and total energy of 1.2μJ/frame.

Acknowledgements:
This work was supported by the Semiconductor Research Corporation (SRC), Princeton Program in Plasma Science and Technology, DARPA and Princeton University. IC fab provided by MOSIS.

References:
[1] Y. Afsar, et al., "Large-Scale Acquisition of Large-Area Sensors Using an Array of Frequency-Hopping ZnO Thin-Film-Transistor Oscillators," *ISSCC Dig. Tech. Papers*, pp. 256-257, Feb. 2017.
[2] T. Moy, et al., "A Flexible EEG Acquisition and Biomarker Extraction System Based on Thin-Film Electronics," *ISSCC Dig. Tech. Papers*, pp. 294-295, Feb. 2016.
[3] B. Glisic, et al., "Strain Sensing Sheets for Structural Health Monitoring Based on Large-area Electronics and Integrated Circuits," *Proc. of IEEE*, vol. 104, no. 8, pp. 1513-1528, June 2016.

 978-1-5386-8531-0/19/$31.00



# Sensors:	N	N	N
# Cycles:	N	√N	(eff. data bandwidth) 1
#LAE/IC Interfaces:	1	2× √N	M
Dynamic Range:	1	1	K
# TFTs:	~10N	N	N×M×log2M

Figure 17.3.1: Summary of CS (M measurements for N sensors), and comparative analysis of systems with traditional hybrid-system-interfacing architectures, as N scales.

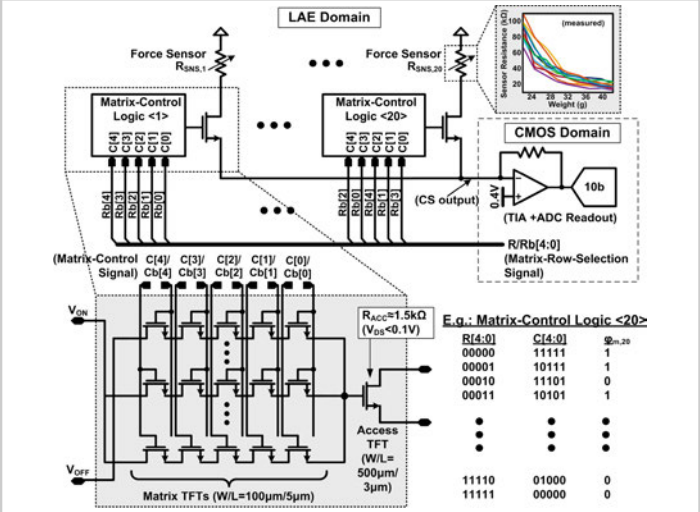


Figure 17.3.2: System for performing readout of sensor-current superpositions, controlled by access and matrix TFTs, using Matrix-Row-Selection signal R/Rb[4:0].

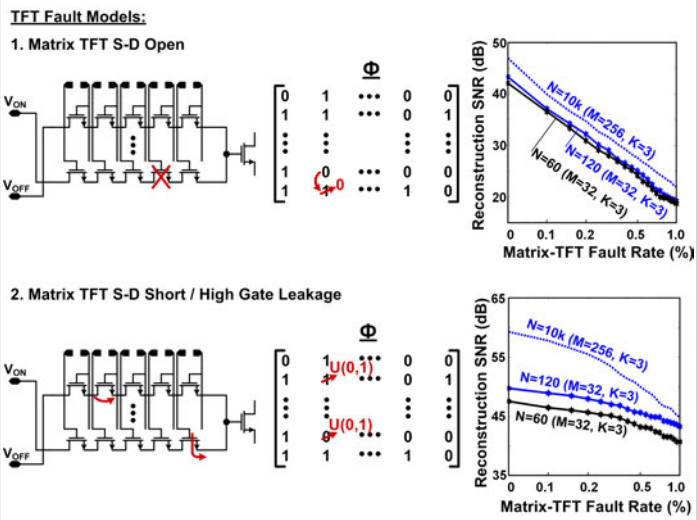


Figure 17.3.3: Analysis of reconstruction-performance tolerance to typical faults in matrix TFTs.

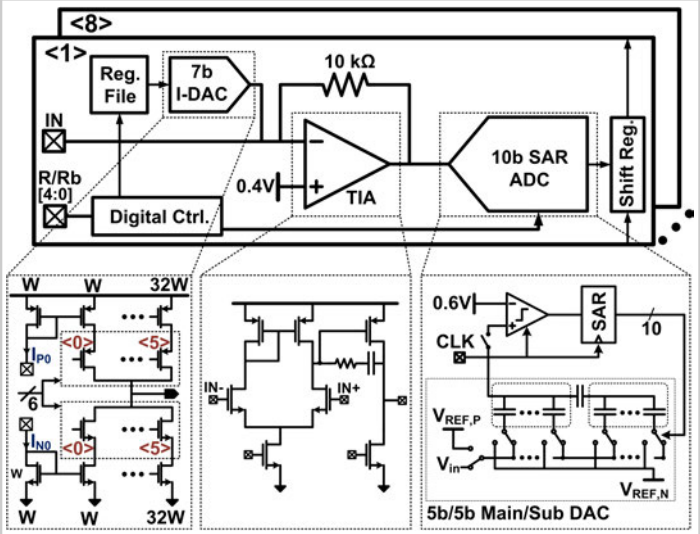
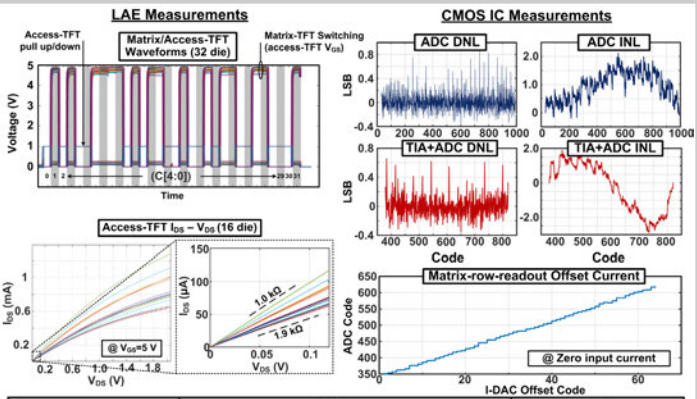


Figure 17.3.4: Block diagram of CMOS IC for matrix-row readout, showing 8 readout channels comprised of offset-canceling I-DAC, TIA, and 10b ADC.



LAE Summary		CMOS Summary				Full-system Summary	
No. sensors	20	Technology	130nm	ADC resolution	10 b	Sensor F range	24 g-40 g
TFT technology	ZnO	VDD	1.2V	ADC energy	212 pJ/conv.	Sensor R range	100 kΩ-15 kΩ
Process temp.	<200°C	Area	4mm²	ADC DNL/INL	0.85/2.0	Sparsity	3
TFTs/sensor	161	No. channels	8	ADC rate	100 kS/s	Recon. error	0.7 kΩRMS
Matrix-TFT speed	1 kHz	TIA power/ch.	37μW	ADC ENOB	8.7 b @50kHz	Frame rate	31 fps
Access-TFT res.	1.5 kΩ	I-DAC res.	6b	TIA+ADC DNL/INL	0.65/2.0	E/frame	1.2 μJ

Figure 17.3.5: Matrix-Control block and access TFTs (left), CMOS ADC DNL/INL, full-system (ADC+TIA) DNL/INL, and I-DAC transfer function (right), with summary table.

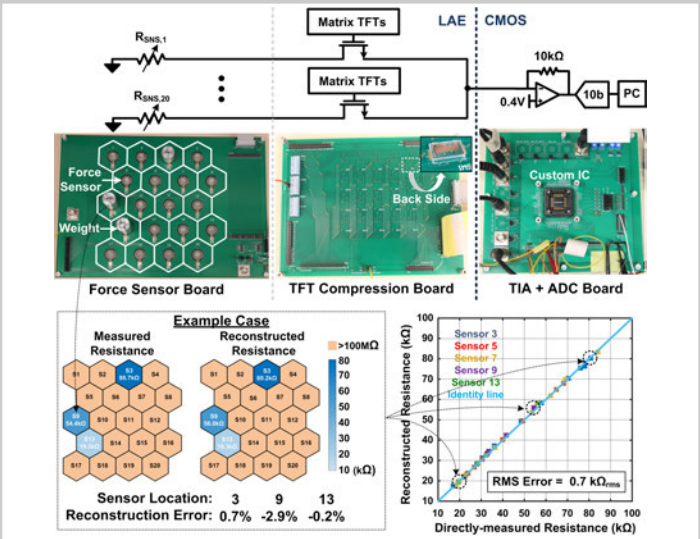


Figure 17.3.6: Measurement and demonstration setup as well as characterization of reconstruction accuracy.

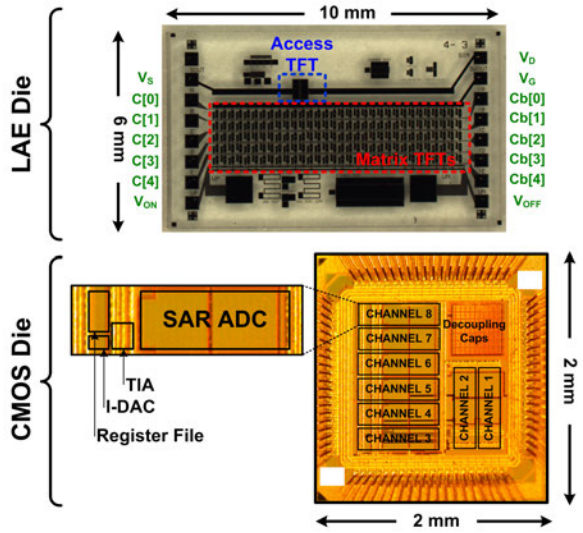


Figure 17.3.7: TFT die micrograph (showing matrix/access TFTs) and CMOS IC die micrograph.

978-1-5386-8531-0/19/$31.00

17.4 16MHz FRAM Micro-Controller with a Low-Cost Sub-1µA Embedded Piezo-Electric Strain Sensor for ULP Motion Detection

Sudhanshu Khanna[1], Michael Zwerg[1], Brian Elies[1], Juergen Luebbe[1], Nagaraj Krishnasawamy[1], Hadi Najar[2], Suman Bellary[2], Wei-Yan Shih[1], Scott Summerfelt[1], Steven Bartling[1]

[1]Texas Instruments, Dallas, TX
[2]Texas Instruments, Santa Clara, CA

Ultra-low Power Microcontrollers (MCUs) [1-4] have played a central role in embedded IoT systems providing programmability, analog and digital processing and control, A/D interfaces, and power management. As IoT applications expand, efficient sensing is increasingly becoming part of MCUs. In this paper we present a 130nm 16MHz Ferro-electric RAM (FRAM) based MCU with a sub-1uA embedded piezo-electric strain sensor and AFE for ULP motion detection (Fig. 17.4.1). To our knowledge, this is the first reported MCU with an embedded motion detection strain sensor. Existing applications that would benefit from such a MCU are applications like toys and remote controls that can turn off while not in use. Motion detection in a key fob improves security by preventing a "man in the middle" attack while the key fob lies stationary at home. Tamper detection and strain gauges are other potential applications. When used in "wake-on-motion" applications, the sensor IP is powered-on at all times waiting for a motion event. Hence minimizing its power consumption is crucial. Also, considering ULP MCUs have many cost-sensitive applications, the sensor must be small in area, and not require any additional masks or special processing steps. A single-chip solution allows reuse of power management, programmability and control circuits already existing in the MCU for use in the strain sensor IP, reducing the system level cost vs a 2-chip solution.

The strain sensor in this work is built using thin film Lead Zirconium Titnate (PZT), which was observed to have a high piezoelectric sensitivity. Other CMOS compatible elements like NWELL resistors also showed piezoelectric behavior but with much lower sensitivity. FRAM, the main memory in this MCU, also uses PZT capacitors (fecaps) as the non-volatile element. The motion detection strain sensor can be manufactured using the same processing steps and masks as FRAM, resulting in no additional processing cost for the sensor. Further, the sensor and AFE together consume only 0.6mm^2, making this a low cost addition to a 130nm based MCU. The IP consumes just 650nA sampling at 100Hz, enabling long battery life.

Figure 17.4.2 shows the sensor IP architecture. The core of the IP is a fecap-based sensor and a 3-stage AFE. The sensor is an array of capacitors and is thus high impedance. Polarized fecaps contain crystal domains that have a polarized dipole associated with them. When strain is applied to a polarized piezo-electric material, the dipoles are displaced resulting in a net surface charge at the terminals of the fecaps. Parasitic leakage from transistors connected to the PZT capacitor terminals prevents them from maintaining the evoked charge indefinitely, so PZT sensors are only suitable for sensing dynamic strain. Gain stages amplify this surface voltage (100's of uV) before it disappears due to leakage. Surface charge that develops when strain is applied is proportional to the magnitude of the fecap polarization. The FSM is used to polarize each fecap upon power-up, setting up the sensor for motion detection. To connect the FSM to the sensor, we must use ultra-low leakage switches since minimizing leakage is critical in enabling detection of low frequency human movement. Sensor, AFE and clock parameters such as sensor size, gain, bandwidth, sample-hold clock duty cycles, clock frequency and strain detection threshold are software programmable. The AFE is sampled to achieve the low power budget. Human motion is slow, and hence a sampling rate of ~200Hz is sufficient, though the AFE is capable of sampling at rates of 1Hz-to-1 kHz. The LNA is a sampled resistive load PMOS-differential amplifier. LNA bandwidth is configurable using capacitors connected to its output. During the long off-time of the sampling period, the drain, source and body of the LNA input pair must be biased at VREF (sensor DC bias) to prevent transistor leakage from draining away the evoked charge on the high impedance sensor or disturbing the sensor's DC bias. AFE Stages 2-3 are switch-capacitor amps with configurable gain. Note that the LNA is an open loop amplifier and thus does not have well controlled offset and gain. LNA offset correction is implemented by trimming current from its resistive load. A current-DAC provides a flexible level of current trim based on software configuration offset control registers. Offset and gain are initially measured by applying a known input voltage (from the REF block) and then offset and gain control registers are set to achieve zero offset and the required gain (using the stage 2-3 PGAs). This procedure is repeated periodically as temperature can change over time. A DAC-subtractor subtracts a user-defined threshold from the AFE output and provides it to a comparator block. If the sensor output exceeds this threshold, the comparator output triggers a wake-up interrupt for the MCU core.

Figure 17.4.3 shows the sensor circuit design. When strain (defined as ΔL/L) is applied, an individual fecap develops only uVs of signal/ustrain. To generate a voltage large enough to be detectable above the noise floor of the LNA, fecaps must be stacked in series. Note that stacking capacitors reduces the effective capacitance of the sensor output. To avoid large charge sharing losses with the LNA input capacitance, eight stacks of fecaps are connected in parallel increasing the capacitance of the sensor output (resulting in an 8 wide x "N" tall fecap stack). A discussion of parasitic effects is now due in order to explain why voltage gain rolls off as fecap stacking height increases. In this process technology, the fecap PZT layer is deposited between active/poly and MET1. Thus, parasitic capacitors exist on the fecap from its terminals to metal layers, poly, active and substrate. Transistors connected to the fecap terminals (low leakage switches) also add parasitic drain capacitance. Parasitic capacitance is lowered by not allowing any metal/poly or active above/below the sensor. Effective sensor capacitance decreases as one traverses upwards from the bottom of a stack of fecaps, but this decreasing effective capacitance charge shares with a parasitic capacitance that remains relatively constant at every level in the stack. Thus, stacking fecaps has diminishing returns. In the absence of parasitics, a 64-series 8-parallel single ended sensor would generate the same signal as a 32-series 8-parallel differential (2-leg) sensor while using the same number of total fecaps (512). However, in the presence of parasitics signal from a single ended 64-stack sensor is degraded by 36%. Using a 32-stack differential sensor cuts the stack height from 64 to 32, reducing parasitic loss to 14% while also providing higher supply and reference noise rejection. Low leakage FSM-Sensor interface switches are built using isolated p-well high-voltage thick-gate-oxide NMOS. By biasing the p-well and source to the same DC voltage as the sensor (drain) we can drastically reduce reverse diode and sub-threshold leakage on the fecap terminals allowing signal detection up-to 1Hz.

Motion in any application produces strain on the PCB that is transferred to the sensor through the package. Thus, the strain sensor characteristics as observed by the integrated sensor are also a function of the PCB mechanical system characteristics (resonant frequency, strain transfer efficiency, etc.) in addition to piezoelectric material properties. Our measurement/evaluation-module (EVM) setup uses conventional PCB material and thickness (FR4, 2 layer, 31 mils), conventional die thickness and an off-the-shelf QFN package. The setup is extremely sensitive and detects very light taps/knocks (picture in Fig. 17.4.6), walking or waving. Figure 17.4.4 shows the sensor output for different motion types along with accelerometer output (located on the same PCB).

Figure 17.4.5 shows the sensor o/p and also shows how the output is modulated by PCB characteristics. The lower frequency range (before PCB resonance) is used to show the PZT sensor's relative temperature independence. Sensitivity is 80mV/g at 60× gain (1.33mV/g at input). Reducing PCB/package/die thickness or designing the PCB resonance to match the target motion increases the sensor output even further. For example, a second smaller and thinner (15mil) PCB with just supply, ground and LED pins connectors designed as an add-on to a primary system board helped achieve even higher sensitivity. Measured total noise is 800uV (input referred, dominated by sensor), and varies <15% across -40C to 85C. Measured sensor IP current is 650nA at 100Hz, and 400nA at 33Hz sampling rate (dominated by analog current). The IP is 0.6mm^2 and needs no additional masks or special processing steps resulting in low cost. This IC is the only reported MCU with an embedded motion detection strain sensor. The closest possible comparisons are accelerometer-MCU SIP/2-chip solutions [5-6], which cost more and/or consume more power due to the use of MEMS and higher power AFEs (essential for high precision applications, but not motion detection). Figure 17.4.6 shows the measurement results and comparison table and Fig. 17.4.7 shows the die micrograph.

References:

[1] J. Kwong, et al., "A 65nm Sub-Vt Microcontroller with Integrated SRAM and Switched-Capacitor DC-DC Converter," *ISSCC Dig. Tech. Papers*, pp. 317-318, Feb. 2008.
[2] M. Zwerg, et al., "An 82µA/MHz microcontroller with embedded FeRAM for energy-harvesting applications," *ISSCC Dig. Tech. Papers*, pp. 333-334, Feb. 2011.
[3] Y. Lee, et al., "A Modular 1mm^3 Die-Stacked Sensing Platform With Optical Communication and Multi-Modal Energy Harvesting," *ISSCC Dig. Tech. Papers*, pp. 402-403, Feb. 2012.
[4] A. Klinefelter, et al., "A 6.45 µW Self-Powered IoT SoC with Integrated Energy-Harvesting Power Management and ULP Asymmetric Radios," *ISSCC Dig. Tech. Papers*, pp. 384-385, Feb. 2015.
[5] ST Microelectronics LIS331DLH Datasheet, https://www.st.com/resource/en/datasheet/lis331dlh.pdf
[6] Analog Devices ADXL362 Datasheet, https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL362.pdf

 978-1-5386-8531-0/19/$31.00

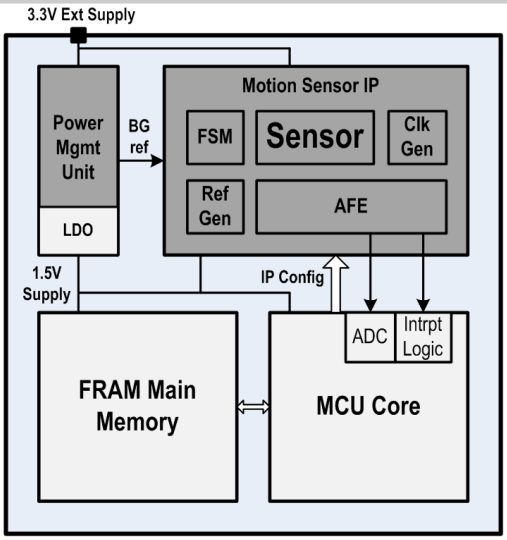


Figure 17.4.1: MCU SoC architecture. The SoC features the sensor IP, power management, memory, and digital and analog peripherals.

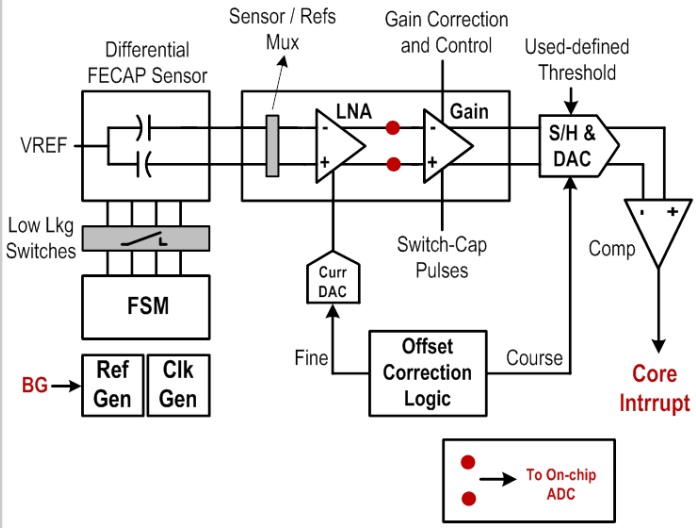
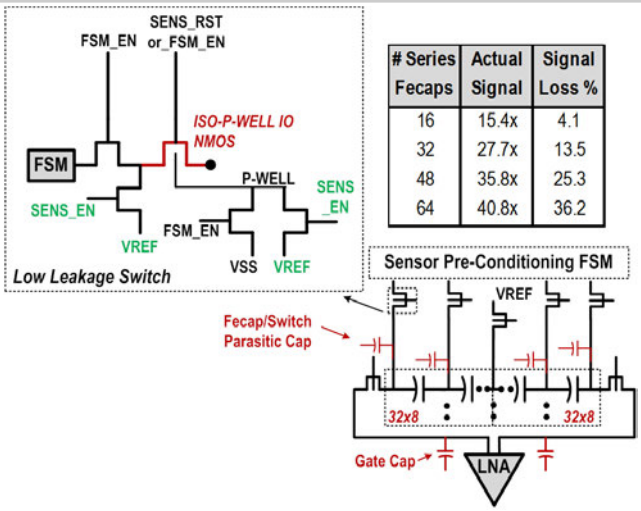


Figure 17.4.2: Sensor IP architecture.



# Series Fecaps	Actual Signal	Signal Loss %
16	15.4x	4.1
32	27.7x	13.5
48	35.8x	25.3
64	40.8x	36.2

Figure 17.4.3: Sensor and switch circuit design. The switch is used in three modes: FSM, sensor, and sensor reset. The table shows how signal loss due to parasitics increases with the sensor series stack height.

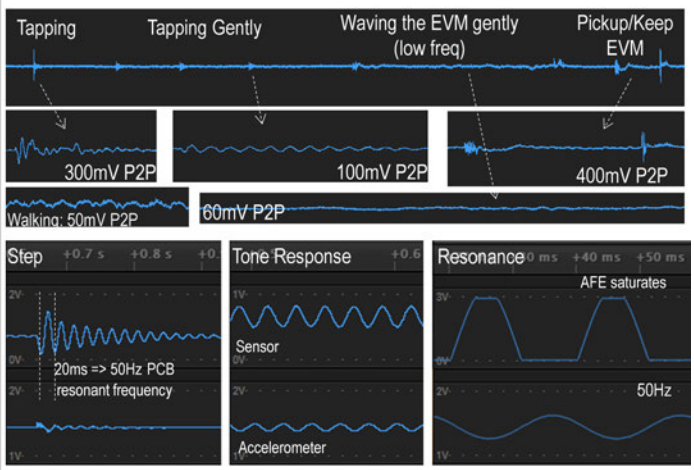


Figure 17.4.4: Measured sensor response to human motion and shaker table based step and sine functions.

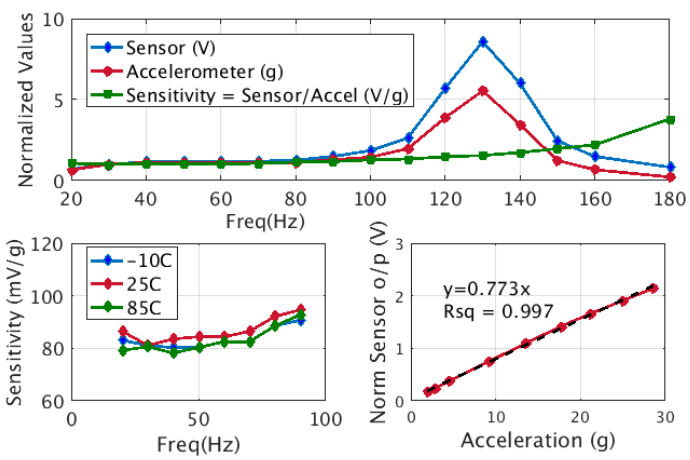


Figure 17.4.5: Measured sensor o/p, accel o/p and sensitivity vs freq and temp (shaker table driven at constant amplitude) and sensor o/p vs amplitude (linearity). Freq<15Hz limited by shaker table.

Technology	130nm
Supply	2.5V-3.6V
Area (Sensor + AFE)	0.6mm2
Sensor IP Power	
Sampling @ 33Hz	400nA
Sampling @ 100Hz	650nA
Total SoC Standby Power	1.4uA
Sampling Range	1Hz - 1kHz

Sensitivity (i/p ref)*	1.33mV/g
Noise (i/p ref)	
AFE	300uV
Sensor+AFE	800uV
LNA Bandwidth	20kHz
Sensor Bandwidth**	>>200Hz
Resolution	9-bit
Range	30g

* g-values measured by on-board accelerometer. Sensitivity = sensor/accel
** Sensor BW measurement limited by PCB bandwidth.

Metric	This work	[5]	[6]
MEMS processing	No	Yes	Yes
Cost Estimate	Low	Medium	High
Power	0.65uA	10uA	3.7uA
Sampling Rate	100Hz	10Hz	100Hz
Suitable for precision appls?	No	Yes	Yes

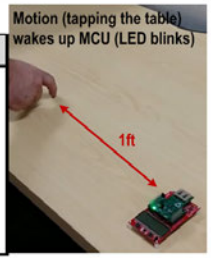


Figure 17.4.6: Measurement results and comparison table. Demo shows MCU waking-up in response to a tap. Power measured at 25C.

17

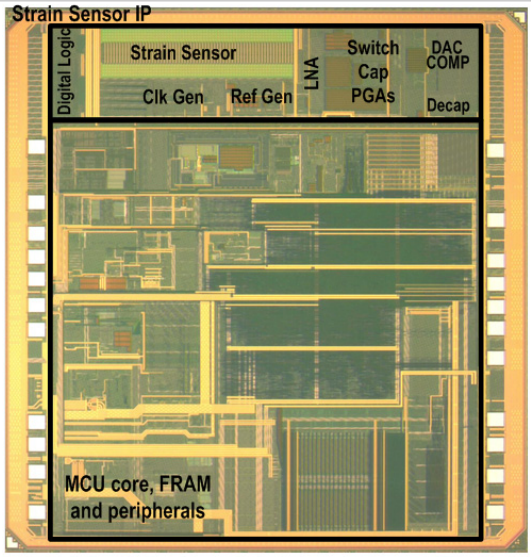


Figure 17.4.7: Die micrograph showing the MCU SoC with the sensor IP.

978-1-5386-8531-0/19/$31.00

17.5 A 0.8mm³ Ultrasonic Implantable Wireless Neural Recording System with Linear AM Backscattering

Mohammad Meraj Ghanbari, David K. Piech, Konlin Shen, Sina Faraji Alamouti, Cem Yalcin, Benjamin C. Johnson, Jose M. Carmena, Michel M. Maharbiz, Rikky Muller

University of California, Berkeley, CA

Miniaturization of implantable neural recording systems to micron-scale volumes will enable minimally invasive implantation and alleviate cortical scarring, gliosis, and resulting signal degradation. Ultrasound (US) power transmission has been demonstrated to have high efficiency and low tissue attenuation for mm-scale implants at depth in tissue [1,2,3], but has not been demonstrated with precision recording circuitry. We present an US implantable wireless neural recording system scaled to 0.8mm³, verified to safely operate at 5cm depth with state of the art neural recording performance an average circuit power dissipation of 13µW, and 28.8µW including power conversion efficiency. Sub-mm scale is achieved through single-link power and communication on a single piezocrystal (Lead Zirconate Titanate, PZT) utilizing linear analog backscattering, small die area, and eliminating all other off-chip components.

A simplified block diagram of the system is shown in Fig. 17.5.1, which consists of an external interrogator and a free-floating mote comprised of only a 0.42mm³ PZT, a 0.25mm² IC, and a flex PCB interposer. The mote is encapsulated with a 10µm layer of parylene-C. The interrogator (Cephasonics, cQuest Cicada) sends a pulse in transmit (TX) mode, which arrives at the mote after one time of flight (ToF). The interrogator then switches to receive (RX) mode to receive the reflected amplitude-modulated (AM) wave from the mote after another ToF. The RX chain then amplifies, filters, and digitizes the received signal for demodulation. Interrogation is performed in a pulse-echo fashion to relax the dynamic range requirements on the RX chain imposed by TX-to-RX acoustic leakage and eliminate the need for a circulator. Thus, to avoid TX/RX interference, the pulse width is set to <2 ToF. Since implant depth can vary, a mote placed close to the interrogator (e.g. ~1cm) would have 2xToF~13µs, too short for wake-up and robust multi-bit digital modulation in a single pulse. To overcome this limitation, the IC performs analog AM backscattering which carries higher information per cycle. The analog signal may be transmitted even if only a few cycles are available per pulse. The addition of a unique subcarrier frequency in each mote can enable simultaneous multi-mote interrogation.

The PZT starts harvesting energy upon the arrival of an incident US wave. To eliminate the need for a large off-chip storage capacitor, the mote wakes up when each pulse arrives and sleeps in absence of a pulse. AC voltage generated across the PZT is rectified and regulated by the IC. At its series resonance frequency (f_r=1.78MHz), the PZT can be modeled as an AC source (V_s) and a series resistance (R_s~4kΩ). The acoustic reflection coefficient (Γ) at the boundary of the mote PZT and tissue can be modulated by tuning the impedance in shunt with the PZT (R_m), as shown in Fig. 17.5.2. However, $\Gamma \propto R_m/(R_m+R_s)$ and is therefore nonlinear with respect to R_m making this method impractical for analog backscattering. Linear backscatter modulation is possible as long as the peak voltage across the piezo terminals (V_{PZ}) is modulated linearly. This can be achieved by connecting synchronously upconverted baseband message current (I_m) to the PZT terminals (Fig. 17.5.2). To save area and complexity, the upconversion is implemented by reusing the available active rectifier on-chip, which automatically synchronizes the conduction period to f_r. During each half cycle, I_m generates a voltage drop across R_s linearly modulating $V_{PZ}=V_{PZ}^+-V_{PZ}^-$. The modulation depth of V_{PZ} is designed such that the rectified voltage is always higher than 1.1V, allowing the LDO to continuously regulate the 1V supply ($V_{DD,1V}$). A small on-chip capacitor (C_{store}=130pF) reduces voltage ripple on the rectified voltage (V_{RECT}). At the highest signal frequency of interest (55kHz subcarrier+5kHz neural signal), the impedance of C_{store} is ~5× larger than the impedance of the PZT, which results in I_m being supplied by the PZT rather than C_{store}. Therefore, in steady state, when the active rectifier pass transistor turns on, the peak voltage across V_{PZ} is, to first order, linearly modulated by I_m.

The IC block diagram is shown in Fig. 17.5.3. As V_{RECT} rises, a power-on reset (POR) is triggered and $V_{DD,1V}$ is generated within 10 cycles (~5µs) after the arrival of the incident wave. Control and timing signals are synchronized by a clock, which is extracted from the PZT voltage ($T_{CK}=1/f_r$). The low-noise amplifier is implemented by an auto-zeroed current-reuse capacitive feedback OTA (Gain 24dB, 250kHz BW required for passing a chopper/subcarrier (~55kHz)). Fast reset switches (ϕ_1, ϕ_2, and ϕ_3) and a power-gated low impedance voltage source are used to quickly charge the input terminals of the OTA to V_{CM}=0.5V in $4T_{CK}$ at the start of the sample. Chopper switches upmodulate the signal to simultaneously reduce 1/f noise of the amplifier and to act as a subcarrier frequency for communication. The kT/C noise sampled on C_s by the reset switches is modulated out of band when the subcarrier is downmodulated on the Rx side in digital. The chopper/subcarrier frequency (f_{CH}) is divided down from the main clock, while its phase is maintained from sample-to-sample by a reset triggered by the POR signal. The amplifier initialization completes in $12T_{CK}$ and the output of the AFE is valid after $22T_{CK}$ (~11µs). 11µs of wakeup time allows backscatter modulation for ~83% of the power-on time at a depth of 5cm. The LNA differential output is forced across a PMOS in triode by means of super source followers generating a linear current over ±200mV of LNA output voltage range. In absence of incident US pulse, the mote PZT is no longer actuated and the circuit turns off gradually as C_{store} discharges (in <10µs), and POR returns to the off state.

The IC was fabricated in a 65nm LP CMOS process. Figure 17.5.4 shows five measured interrogation events of modulated V_{PZ} in response to a $20mV_{PP}$ input sine wave. The 11µs start-up time and subcarrier modulation at f_{CH} is observable on each pulse. The demodulated output signal and its corresponding spectrum are plotted for a 300Hz, $20mV_{PP}$ input sine wave whose measured SFDR and THD are -50dB and -44dB, respectively. No harmonic tones are observed in the spectrum for a $10mV_{PP}$ signal.

The measured static input-output characteristic is shown in Fig. 17.5.5 (output defined as the envelope of V_{PZ}). A 23dB voltage gain over $20mV_{PP}$ of input range with <1.2% static nonlinearity is achieved. PEDOT:PSS coated electrodes (200x200µm², electroless nickel immersion gold plating) present <10mV DC offset, which is within the V_{IN} range. The demodulated input-referred noise spectral density with and without subcarrier modulation is plotted in Fig. 17.5.5. The total input-referred noise spectral density is 328nV/√Hz, and is dominated by carrier noise, which contributes 319nV/√Hz, therefore the noise contributed by the chip is 76nV/√Hz.

An *in vitro* measurement is shown in Fig. 17.5.6, where the assembled mote is suspended at a distance of 5cm away from the interrogator in water. Backscatter modulation and demodulated signal are shown for a 300Hz, $20mV_{pp}$ sine wave applied at the input. A comparison with recently published fully integrated sub-mm³ neural recording implants is shown in the table in Fig. 17.5.7. We introduce a linearization technique for AM backscattering in ultrasonic implants. This design occupies the same volume as the smallest neural recording implant, but improves performance by combining low-noise, power-efficient, and linear neural recording with a PEF 2× lower, a depth 2.5×, and a static nonlinearity 40× lower than state of the art [1,4,5].

Acknowledgments:
The authors thank DARPA BTO, the sponsors of BWRC and TSMC for chip fabrication. Thanks to Ka Yiu Lee and Burak Eminoglu for technical discussion.

References:
[1] D. Seo, et al., "Wireless recording in the peripheral nervous system with ultrasonic neural dust". *Neuron*, vol. 91, pp. 529-539, 2016.
[2] T. C. Chang, et al., "27.7 A 30.5mm3 fully packaged implantable device with duplex ultrasonic data and power links achieving 95kb/s with <10E–4 BER at 8.5cm depth," IEEE *ISSCC Dig. Tech. Papers*, pp. 460-461, Feb. 2017.
[3] B. C. Johnson, et al., "StimDust: A 6.5mm3, wireless ultrasonic peripheral nerve stimulator with 82% peak chip efficiency," *IEEE CICC*, pp. 1-4, 2018.
[4] S. Lee, et al., "A 330µm×90µm opto-electronically integrated wireless system-on-chip for recording of neural activities," *ISSCC Dig. Tech. Papers*, pp. 292-294, Feb. 2018.
[5] P. Yeon, et al., "Towards a 1.1 mm2 free-floating wireless implantable neural recording SoC," *IEEE CICC*, pp. 1-4, 2018.

 978-1-5386-8531-0/19/$31.00

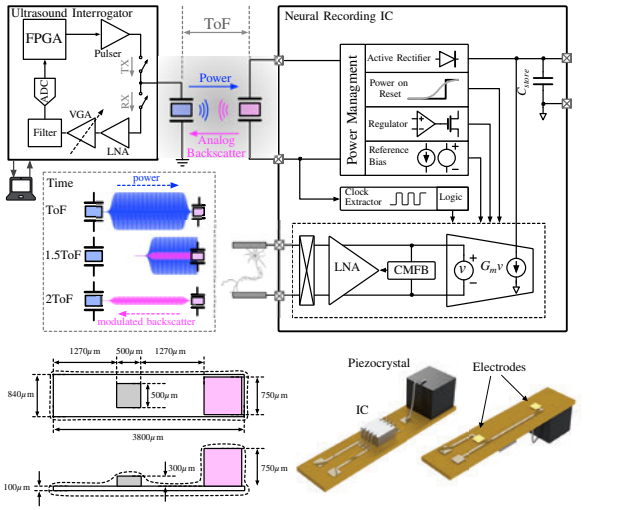


Figure 17.5.1: Simplified system block diagram, time-space conceptual diagram of power transmission and modulated backscattered waves. Assembled mote diagram.

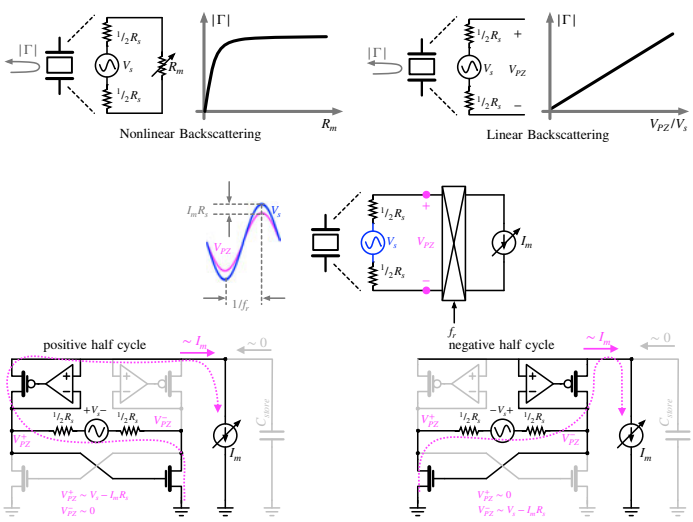


Figure 17.5.2: Backscatter uplink: nonlinear Γ vs. Rm. Linear Γ vs. V_{pz}. Reusing available active rectifier as a synchronous base-band I_m up-converter.

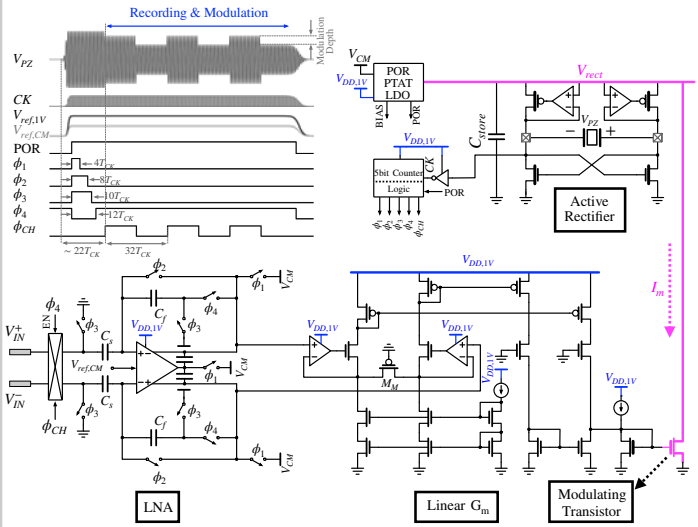


Figure 17.5.3: IC circuit schematics and timing diagram for a single sample pulse. Front-end LNA, subcarrier generation, and linear modulation are shown.

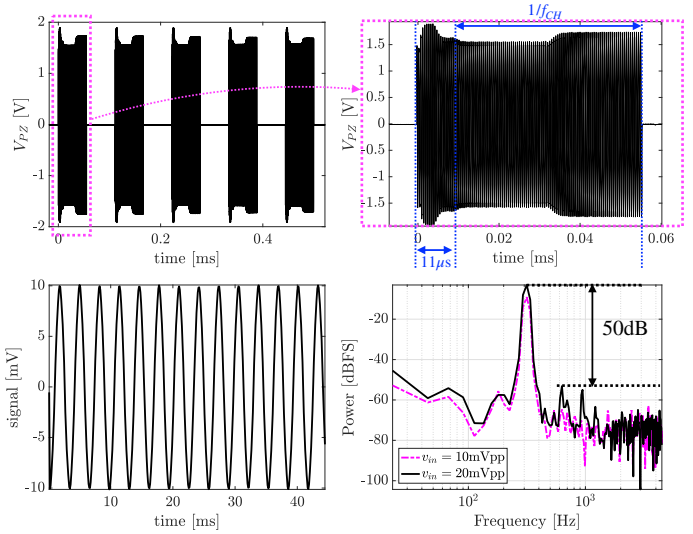


Figure 17.5.4: Five consecutive sample pulses for a 300Hz, 20mVpp input signal, the demodulated input-referred signal, and its corresponding spectrum.

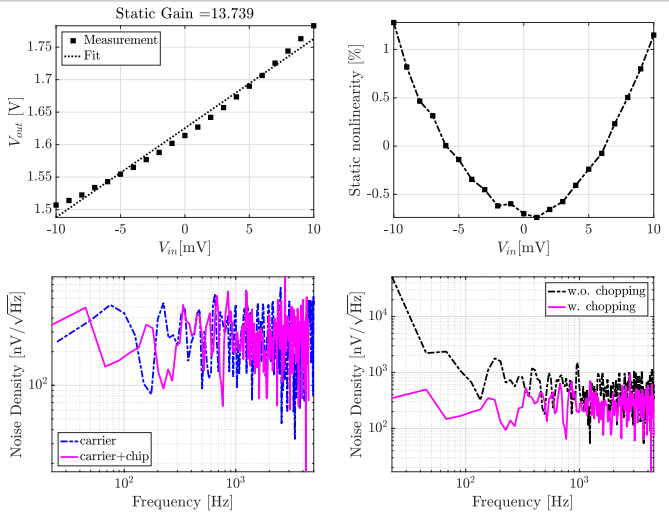


Figure 17.5.5: Static gain and nonlinearity (top); input-referred noise compared with the noise of the carrier alone, with and without chopping/subcarrier (bottom).

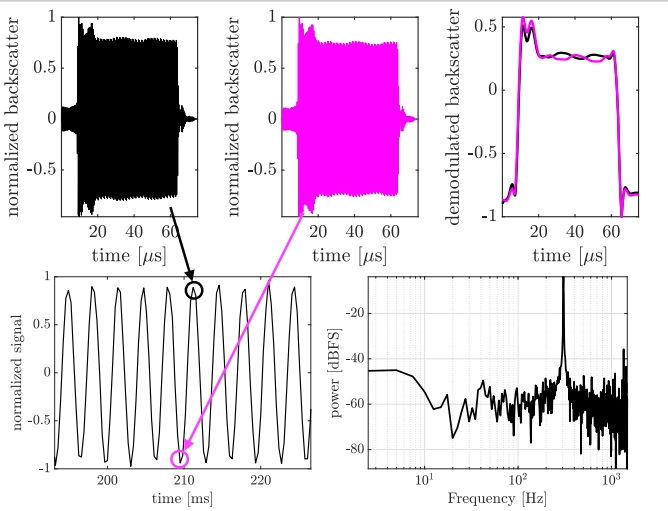
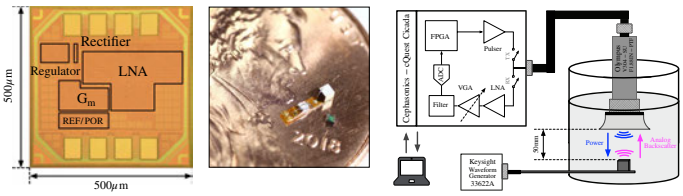


Figure 17.5.6: In vitro measurement results at 5cm depth: two individual backscatter signals, their corresponding AM demodulation and reconstructed 300Hz signal.



		Neuron'16	ISSCC'18	CICC'18	This work
Link		Ultrasound	Optical	RF	Ultrasound
Back telemetry		AM Backscatter	Analog PPM	IR-UWB	AM Backscatter
Implant depth (mm)		8.8	–	20	50
Total Volume (mm^3)		0.8	–	<1	0.8
Fully packaged		Yes	–	Yes	Yes
Depth/Volume (mm^2)		11	–	30.3	61.7
Technology (nm)		65	180	350	65
Total power (μW)		–	0.9	296*	28.8
IC area (mm^2)		500×450	90×330	1050×1050	500×500
Wake-up time (μs)		3.3	1000	110	11
THD (dB) @10mV_{pp}/ @20mV_{pp}		–	–	–	-52/-44
Static nonlinearity (%)		–	49†@6mV_{PP}	–	1.2@20mV_{PP}
LNA	noise floor (μV_{rms})	180	42	3.78	5.3
	Bandwidth (kHz)	5	10	12	5
	power (μW)	–	0.52	42.9	4
	NEF/ PEF	–	12.3/136	6.6/78.4	5.87/34.55
	Gain (dB)	0	24	53	24

†estimated
*after AC-to-DC

Figure 17.5.7: Chip and assembled mote micrograph. In vitro measurement setup and performance comparison table.

978-1-5386-8531-0/19/$31.00 ©2019 IEEE

17.6 A Sub-40μW 5Mb/s Magnetic Human Body Communication Transceiver Demonstrating Trans-Body Delivery of High-Fidelity Audio to a Wearable In-Ear Headphone

Jiwoong Park, Patrick P. Mercier

University of California, San Diego, La Jolla, CA

Emerging wearable devices such as wireless headphones, smart glasses, and medical monitors require increasingly high-throughput wireless communications at ultra-low-power. Since far-field RF has significant path loss around the human body (e.g., up to 70dB at 2.4GHz), most RF body-area-network (BAN) systems such as Bluetooth Low Energy (BLE) have significant energy-expensive amplification requirements consuming milliwatts of power, and thus do not meet the energy demands of emerging small devices. Popular wireless earbuds, for example, only achieve a battery life of a few hours. Exasperatingly, emerging high-fidelity streaming audio and video content requires higher data rates than what BLE can currently accommodate. Human body communication (HBC) systems, for example ones based on electric fields (eHBC), in theory have lower path loss and can thus potentially offer more efficient links [1-2]. However, measurements from form-factor-accurate prototypes reveal path loss that is still rather large (e.g., 30-to-45dB across 20cm [3]), with unfortunately severe variation with posture and environments that requires energy-expensive compensation. Since the human body is magnetically inert, magnetic HBC (mHBC) systems, illustrated in Fig. 17.6.1, offer much lower path loss (e.g., 5-to-30dB over 1m [4]) without severe variation, and can thus theoretically achieve lower communication energy. However, to date there has not been any mHBC transceiver (TRX) developed to exploit this inherently efficient communication channel, and if there were, the high Q of employed coils would limit data rate of a straightforward approach to <800kb/s.

This paper presents the design of an mHBC TRX that consumes <40μW at a data rate of 5Mb/s, for the lowest energy/bit amongst pragmatic prior art wireless meter-range BAN transceivers. Efficient operation is achieved by: 1) exploiting the intrinsically low path loss between high-Q mHBC coils to reduce TX output power and RX gain/noise requirements; 2) employing an mHBC coil as a high-Q resonant element in an energy-efficient directly-OOK-modulated power oscillator (PO); 3) communicating at a high data rate (5Mb/s) despite the high-Q TX coil, which nominally limits bandwidth to <800kHz, via an all-digital frequency-locked-loop (ADFLL)-based synchronous injection-locked accelerated quenching and kick-start circuit, which also serves to dynamically tune the TX resonant capacitor to maintain center frequency lock between TX and RX while compensating the impedance variation of body-worn coils; and 4) biasing dynamic-threshold MOS (DTMOS) RX amplifiers and envelope detector (ED) in deep subthreshold, all at link-budget-appropriate noise levels.

A block diagram of the proposed mHBC TRX is shown in Fig. 17.6.2. The TRX utilizes a power oscillator to directly generate the 40MHz carrier by using the coil itself as the resonant element in the oscillator [5], as shown in Fig. 17.6.3. However, since the Q of the TX coil can be large (e.g., 50), driving the PO on and off for OOK modulation via the tail transistor cannot be normally accomplished at a data rate faster than 800kb/s. This is one of the reasons why conventional near-field communication (NFC) protocols implementing low-frequency (<13.5MHz) near-field coupling (as opposed to simultaneous near- and far-field coupling at a higher frequency in mHBC [4]) cannot support sufficiently high data rates. Here, to increase data rate, transistors M_{s1} and M_{s2} are deployed to rapidly shunt the inductor current during transmissions of logic '0's. On the other hand, to rapidly kick-start oscillations when transmitting logic '1's, which would normally take more cycles at 40MHz than are available in a 5Mb/s symbol period, injection transistors M_{i1} and M_{i2} are deployed, improving oscillation startup time by 20×. To maximize the effect of injection-locked kick-start, the resonant frequency of the PO, which can change slightly with posture or environmental variation (though not as severely as in eHBC), is precisely set to the same frequency as the crystal-based injection source for synchronous injection via a 4×4 unit-capacitor array that is dynamically tuned via an ADFLL, also shown in Fig. 17.6.3. This also helps to minimize any frequency mismatching between TX and RX. The fast control bit output (2.5Moutput/s) of the ADFLL enables frequency tuning within 10-to-20μs every 1-to-10ms to minimize wait time during calibration.

The RX, shown in Fig. 17.6.3, is designed using a common-source (CS) preamplifier, an active CS envelope detector (ED), a single-ended-to-differential (S-to-D) buffer, and variable gain amplifier (VGA). All RX circuits are biased in subthreshold and utilize DTMOS inputs to improve transconductance or conversion gain by up to 32% iso-current. The ED output drives a S-to-D buffer designed with low-V_t devices to minimize input capacitance to help ensure 2.5MHz bandwidth at the output of the low-current ED; the S-to-D then drives a differential 3b VGA. The VGA output is bit-sliced via a regenerative comparator at the appropriate common-mode voltage thanks to the VGA's common-mode feedback (CMFB) circuit.

The mHBC transceiver is fabricated in 0.12mm^2 of core area in 65nm CMOS. Figure 17.6.4 (top left) shows the output waveform of the TX under worst-case modulation of alternating 1s and 0s at 5Mb/s. Thanks to the accelerated quenching and kick-start circuit, the ASK-like modulation index improves from 15.4% to 84.7% compared to a conventional power oscillator, thereby enabling the desired high data rate in a manner that is compatible with the low-complexity RX. The spectrum of a signal received by an mHBC coil across a 10cm channel is shown in Fig. 17.6.4 (top right). Here, it can be seen that the ADFLL automatically re-tunes the intentionally-mis-tuned PO resonance frequency (~40.4MHz) of TX to the correct carrier frequency (40MHz), while also demonstrating that injection locking further improves total TX output power by 17.4dB while providing the rapid kick-start behavior described above. The inset plot illustrates the frequency-tracking capability of the TX, while validating the fast frequency settling (<10μs) of the ADFLL. BER testing in Fig. 17.6.4 (bottom left) under asynchronous operating conditions with 2× oversampling demonstrates receiver sensitivities of -63.5dBm to -56dBm at data rates from 1.25-to-5Mb/s, respectively, and from -57dBm to -54dBm according to VGA gain setting at 5Mb/s. When operating over a distance of ~1m across a human body, a pair of mHBC TRXs, with -24.8dBm TX output power and the -56dBm RX sensitivity setting, achieved a TX-to-RX latency at 5Mb/s of 388ns with error-free operation, as shown in Fig. 17.6.4 (bottom right).

To demonstrate feasibility of a practical application, two-channel analog audio data was generated by a smartphone and delivered via a two-channel 24b ADC and AES3 encoder to an mHBC TX employing a 2-turn 1mm width coil printed on the outline of 11×5.5cm^2 1oz copper PCB mounted on the smartphone. An mHBC RX, employing a 32mm-diameter coil mounted within an in-ear headphone prototype, wirelessly received data, playing audio data through a speaker via an AES3 decoder and DAC. A block diagram of the test setup is shown in Fig. 17.6.2 (bottom), a photograph of the test in progress is shown in Fig. 17.6.5, and the developed components are shown in Fig. 17.6.6 (top). Here, audio is shown to be delivered from the smartphone case to the headphone prototype across the body with no bit errors and <90μs latency (dominated by the 88μs of latency of the AES3 decoder).

As summarized in Fig. 17.6.6 (bottom), the TX consumes 18.6μW at -24.8dBm of output power, resulting in an efficiency of 3.7pJ/bit at 5Mb/s. At -56dBm sensitivity, the RX consumes 6.3μW for an efficiency of 1.3pJ/bit. When including the power of a 17.2μW crystal oscillator, the TX and RX consume 35.8μW (7.2pJ/bit) and 23.5μW (4.7pJ/bit), respectively, all while reliably covering a ~1.5m BAN. This is the lowest power and most efficient TRX amongst prior-art that reliably communicate over a BAN and that includes the power consumption of frequency synthesis. A die micrograph is shown in Fig. 17.6.7.

Acknowledgement:
This work was supported in part by the National Science Foundation under Grant No. 1751293.

References:

[1] S. Maity, et al., "A 6.3pJ/b 30Mbps -30dB SIR-tolerant Broadband Interference-Robust Human Body Communication Transceiver using Time Domain Signal-Interference Separation," *IEEE CICC*, pp. 1–4, Apr. 2018.

[2] J. Jang, et al., "4-Camera VGA-resolution capsule endoscope with 80Mb/s body-channel communication transceiver and Sub-cm range capsule localization," *ISSCC Dig. Tech. Papers*, pp. 282–283, Feb. 2018.

[3] J. Park, et al., "Channel Modeling of Miniaturized Battery-Powered Capacitive Human Body Communication Systems," *IEEE. Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 452–462, Feb. 2017.

[4] J. Park, et al., "Magnetic Human Body Communication," *IEEE EMBC*, pp. 1841–1844, Aug. 2015.

[5] P. P. Mercier, et al., "A Sub-nW 2.4 GHz Transmitter for Low Data-Rate Sensing Applications," *IEEE J. Solid-State Circuits*, vol. 49, no. 7, pp.1463-1474, July 2014.

 978-1-5386-8531-0/19/$31.00

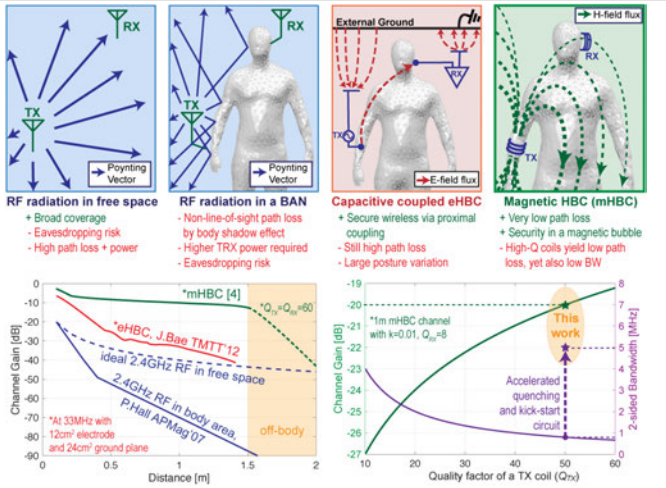


Figure 17.6.1: The path loss of far-field RF, eHBC, and mHBC compared (top and bottom left); increasing the Q of the TX coil decreases path loss, but also nominally decreases bandwidth without the proposed accelerated quenching and kick-start circuit (bottom right).

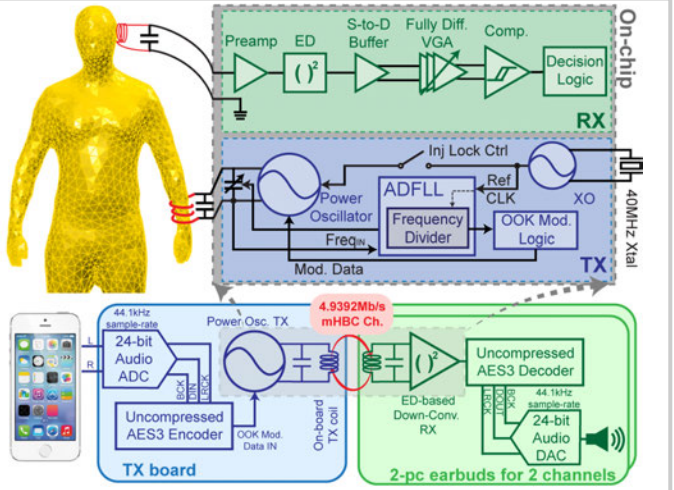


Figure 17.6.2: Block diagram of proposed mHBC transceiver (top); block diagram of the audio streaming demonstration test setup delivering two channels of uncompressed 44.1kHz 24b audio to demonstrate the high-throughput capabilities of the mHBC link (bottom).

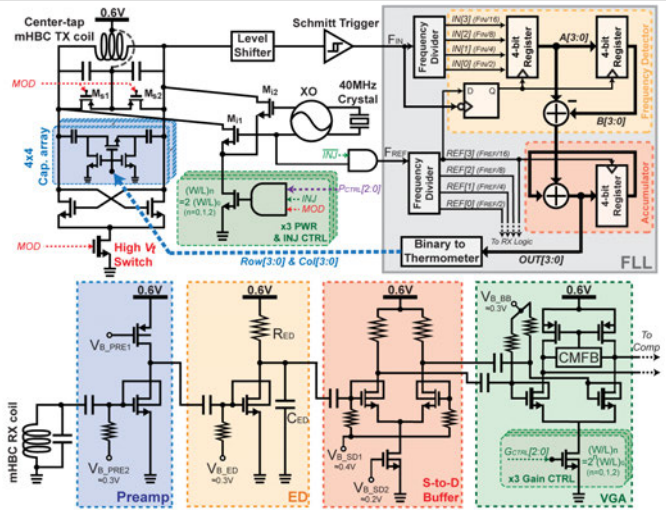


Figure 17.6.3: Schematics of the TX power oscillator and injection locking circuit for the faster-than-Q data-rate modulation (top); the sub-Vt DTMOS envelope-detector based RX (bottom).

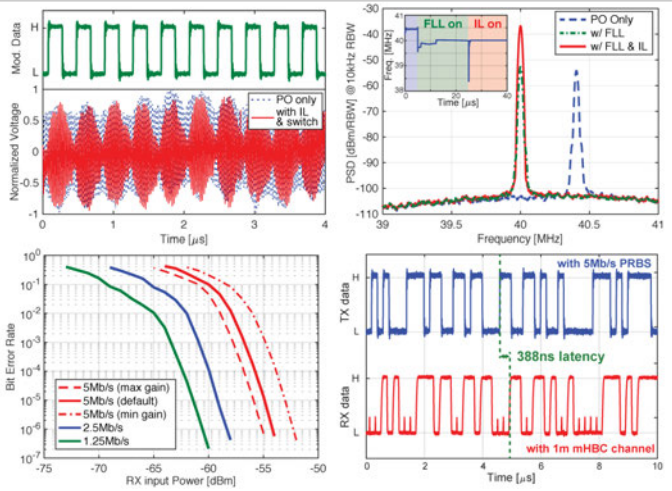


Figure 17.6.4: Measured 5.5× improved worst-case OOK-modulated TX output waveform (top-left), measured spectrum and frequency transient (inset) of dynamic resonance tuning by the ADFLL (top-right), measured RX BER curves (bottom-left), and measured cross-body waveform results (bottom-right).

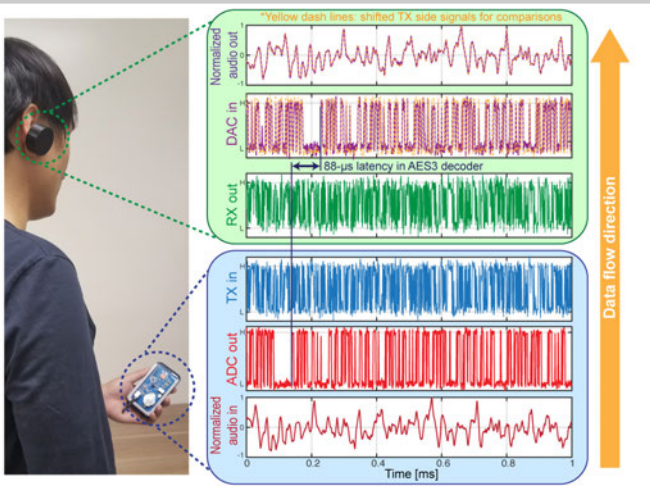


Figure 17.6.5: System demonstration results of high-fidelity audio data streaming (2-channel 24b 44.1kHz audio) via an mHBC link supporting 4.9392Mb/s AES3 format to demonstrate the high-throughput capabilities of the mHBC TRX.

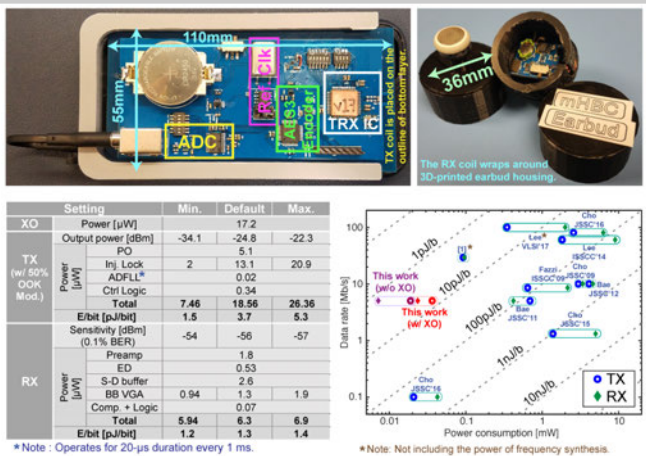


Figure 17.6.6: Photograph of the battery-operated audio streaming prototypes: the TX unit fits within a smartphone case (top-left) and the earbud-based RX prototypes were 3D printed for demonstration purposes (top-right); table of specifications (bottom-left), and the state-of-the-art comparison chart (bottom-right).

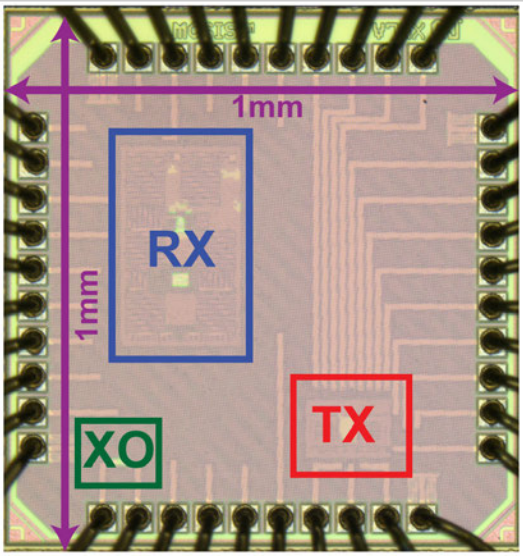


Figure 17.6.7: Die micrograph of the proposed mHBC transceiver.

978-1-5386-8531-0/19/$31.00

17.7 A 7.0fps Optical and Electrical Dual Tomographic Imaging SoC for Skin-Disease Diagnosis System

Yongsu Lee, Kwantae Kim, Jiwon Lee, Kyoung-Rog Lee, Surin Gweon, Minseo Kim, Hoi-Jun Yoo

KAIST, Daejeon, Korea

Recently, skin disease patients have been rising rapidly, and more than 5.4 million cases of skin cancer were treated in over 3.3 million people in the U.S. in 2012, and more people are diagnosed with skin cancer each year than all other cancers combined in the U.S. [1]. Because a malignant lesion is difficult to distinguish from pigmented nevus, elaborate detection devices are needed for accurate diagnosis [2]. Optical Coherence Tomography and Laminar Optical Tomography are mainly used for skin disease diagnosis, however, equipment for those tomography are too bulky to be used in house. The Electrical Impedance Spectroscopy is also applied to analyze skin disorders [3], but its equipment requires assistance from experts.

In this paper, a coin-sized light-weight (3g) and portable system that simultaneously takes optical and electrical domain skin images is proposed for accurate skin disease diagnosis. It uses Diffused Optical Tomography (DOT) and Electrical Impedance Tomography (EIT) concurrently for hemodynamic and electrical characteristics measurements of the skin.

Figure 17.7.1 shows the proposed skin disease diagnosis system. The circular system has only a 29mm diameter and 3g weight enabling a seamless measurement on any part of the human body. The system is composed of two modules: Micro-sensor module and chip module. Micro-sensor module is implemented on flexible PCB substrate containing 19 micro-cells. The micro-cell consists of 3 micro-LEDs with different wavelengths (622, 528, and 470 nm), a photodiode (PD) for DOT, and a surrounding electrode for EIT. Three wavelengths are selected because the number of main absorbers in human skin is three with different absorption peak wavelengths, which is melanin, Hb, and HbO_2. Hemodynamic characteristics is extracted from DOT for diagnosis of Angioma. Frequency difference EIT (FD-EIT) is adopted for acquiring electrical information of lesion, which can be used for diagnosis of Melanoma.

Figure 17.7.2 shows the overall block diagram of the SoC. It consists of: 1) Power management unit (PMU) for generating stable two voltage domains (7V for V_{LED}, 1.8V for V_{DD}), 2) Current driver for 3 wavelengths CDMA DOT and 2 frequencies combined generator for FD-EIT, 3) DOT receiver with fast settled DC rejection scheme, 4) EIT receiver implemented with weaver demodulator and IQ phase calibrator, and 5) Digital module for data acquisition and communication with external devices.

Figure 17.7.3 describes the dual calibration loop between boost converter and DOT current driver. The switched capacitor type is adopted because higher efficiency can be achieved in <10mW output power range compared to inductive boost converter [4]. The boost converter consists of 5 cells of 1:2 bi-phasic voltage doubler. The switch configuration controlled by 5-bit control signal (S_C), is determined according to comparison results between battery voltage (V_{BAT}) and reference voltage. S_C switches the connection of unit cells to boost V_{LED} by an initial coarse step voltage increase. Then, V_{LED} calibration loop controls the output impedance of boost converter with negative feedback. The VCO at the next stage of analog comparator controls operation frequency (f_{CLK}) of boost converter, and stabilizes V_{LED} to 7V with fine voltage step. To avoid breakdown from high V_{GS} (>5 V), double-step switch is proposed. The (A_0 – 1) V is generated from a charge pump and connected to the V_{SS} terminal of the level shifter circuits, and then V_{GS} of M_0 and M_1 is always < 5V. Due to process mismatch and aging, the value of Extinction Ratio (ER), $\Delta P/\Delta I$, may vary LED by LED. The PD (mPD) from inner micro-cell monitors ER, and 4-bit controlled current steering DAC (I_{EDAC}) increases current level until ΔI equals to I_{ER}.

Figure 17.6.4 shows the circuit schematics and the measurement results of the DOT receiver. To measure accurate hemodynamic characteristics from optical sensing, fast imaging speed (3-to-100 fps) and wide dynamic range (>80 dB) are required. The settling time of analog DC rejection scheme is too long (>1s) to measure constant data, and digital DC rejection scheme has limited dynamic range due to quantized rejection resolution [5]. To achieve short settling time and high resolution, duty-cycled resistor based DC rejection loop is proposed. In the feedback loop, the cut-off frequency is decided with R_E of which value is amplified by 1/D with the help of duty-cycled resistor. The "D" is derived from the ratio of pulse-width to period of clock. When the clock period sweeps from small to large values, the cut-off frequency decreases while reducing the settling time with the same DC rejection. When there is 1μA input current fluctuation, 100 times lower settling time (50μs) is measured compared to the fixed period case as shown in the measurement results. To reduce the effect of parasitic capacitance (C_P) between switching MOSs, PCIS is proposed as shown in the figure. In addition, input capacitance (C_{IN}) cancellation scheme is adopted to increase the dynamic range of TIA. C_{IN} is the main cause of input referred noise, and positive feedback loop with source follower [6] compensates for it. Therefore, 100.4dB dynamic range of TIA is measured with 9.57pA_{rms} [50Hz] input referred noise.

Figure 17.7.5 shows the architecture and schematics of the EIT receiver. The weaver architecture is commonly used as the image rejection scheme in RF circuit [7]. The proposed modified weaver demodulator utilizes the image rejection channel as f_2 channel for FD-EIT. This architecture is effective for narrow bandwidth between f_1 (100kHz) and f_2 (80kHz) of FD-EIT without necessity of high order low-pass filter. In the FD-EIT, since only 1° phase error leads to 2.8% demodulation error, phase detector and phase interpolator are used to compensate for I/Q clock phase mismatch.

Figure 17.7.6 shows the measurement results of an in-vitro and in-vivo environment. 1.2% w/w agar phantom is used instead of human skin. A 5mm diameter of India-ink-mixed agar is modeled as pigment nevus, and 5mm × 5mm square-cut ink-dyed carrot is utilized as a malignant lesion. From the DOT image, two regions are unable to be distinguished due to similar absorbance parameter. However, an EIT image differentiates two regions due to different electrical characteristics with a blurred boundary image. By combining DOT and EIT together, images with both advantages, high sensitivity (0.5mm) and high selectivity (2mm) can be obtained. In addition, even a 3-D image can be reconstructed, and 80% thresholding shows the image with positions of only malignant lesion. Two kinds of letter (N, S) tattoo are tested for the in-vivo measurements, and the DOT image shows a similar shape to the real image. 3-wavelength measurements can extract the change of HbO_2 while pressure is applied to skin.

Figure 17.7.7 shows the chip micrograph, performance summary of SoC, and comparison table between DC rejected optical sensors. The proposed SoC is fabricated in 180nm high voltage CMOS process. The die area is 3.3mm × 5.0mm. The average power consumption is 7.54mW, and 2.8 ~ 4.2 V input voltage convert to 1.8V for V_{DD} and to 7V for V_{LED}. CDMA DOT current driver operates with 16MHz clock frequency and 3.9kHz chip rate to achieve 7.0fps hemodynamics imaging speed. Current range is wide from 0.05mA to 6.4mA. The DOT receiver achieves a fast settling time withq DC rejection (50s) and wide dynamic range (100.4dB). Weaver demodulator based EIT receiver operates with I/Q mismatch calibrator. As a result, 0.5mm sensitivity and 2mm selectivity dual tomographic imaging can be obtained with the proposed SoC for skin disease diagnosis system.

References:

[1] American Cancer Society, "Cancer Facts and Figures 2018."
[2] T. J. Muldoon, et al., "Analysis of skin lesions using laminar optical tomography," *BIOMEDICAL OPTICS EXPRESS*, pp. 1701-1712, June 2012.
[3] J. Malvehy, et al., "Clinical performance of the Nevisense system in cutaneous melanoma detection: an international, multicentre, prospective and blinded clinical trial on efficacy and safety," *British Journal of Dermatology*, pp. 1099-1107, Oct. 2014.
[4] G. Villar-Pique, et al., "Survey and Benchmark of Fully Integrated Switching Power Converters: Switched-Capacitor Versus Inductive Approach," *IEEE Trans. Power Electronics*, vol. 28, no. 9, pp. 4156-4167, Sept. 2013.
[5] E. S. Winokur, et al., "A Low-Power, Dual-Wavelength Photoplethysmogram (PPG) SoC With Static and Time-Varying Interferer Removal," *IEEE TBCAS*, vol. 9, no. 4, pp. 581-589, Aug. 2015.
[6] C.-M. Tsai, "A 40mW 3Gb/s Self-Compensated Differential Transimpedance Amplifier With Enlarged Input Capacitance Tolerance in 0.18μm CMOS Technology," *IEEE JSSC*, vol. 44, no. 10, pp. 2671-2677, Oct. 2009.
[7] S.-C. Hwu and B. Razavi, "An RF Receiver for Intra-Band Carrier Aggregation," *IEEE JSSC*, vol. 50, no. 4, pp. 946-961, Apr. 2015.

 978-1-5386-8531-0/19/$31.00

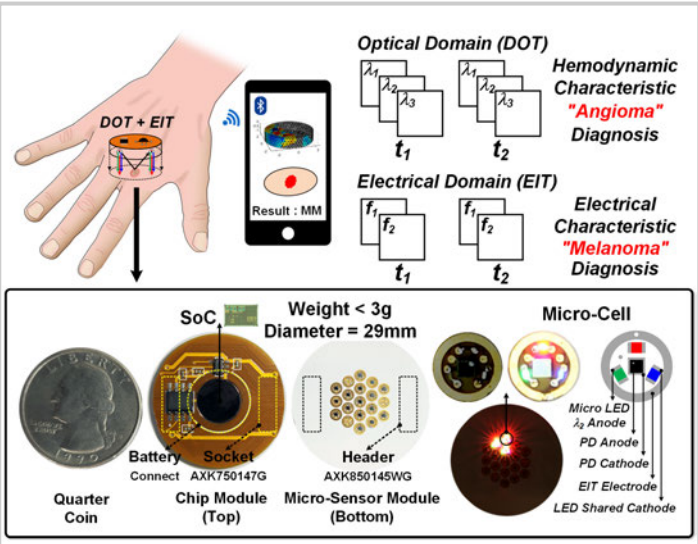


Figure 17.7.1: Proposed skin disease diagnosis system.

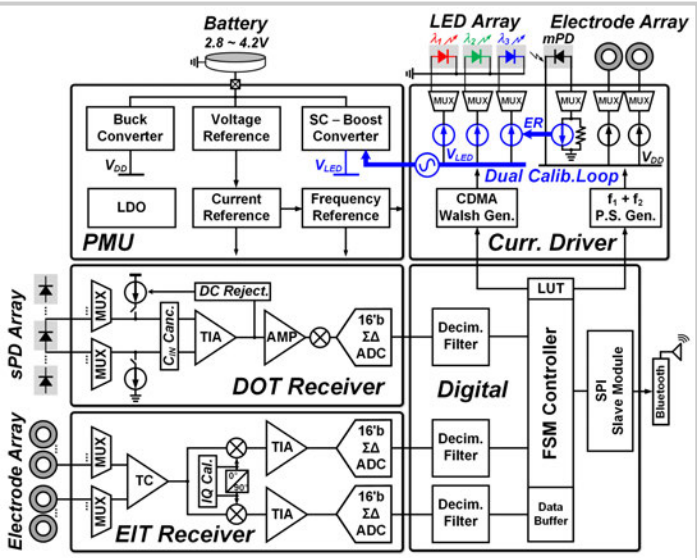


Figure 17.7.2: Overall block diagram of SoC.

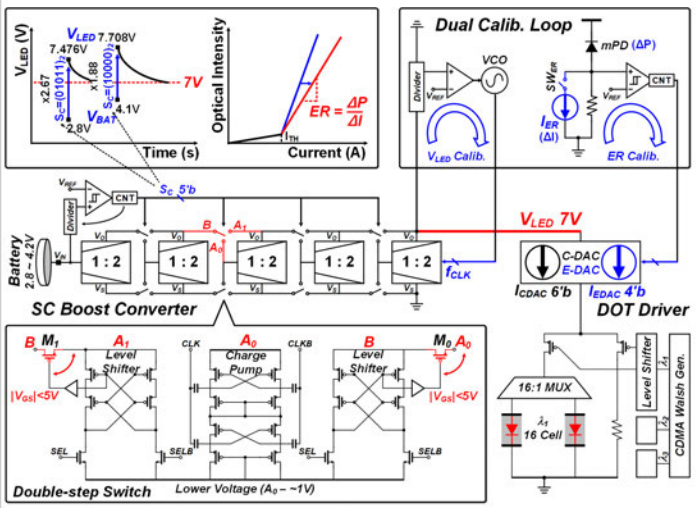


Figure 17.7.3: The dual calibration loop between boost converter and DOT current driver.

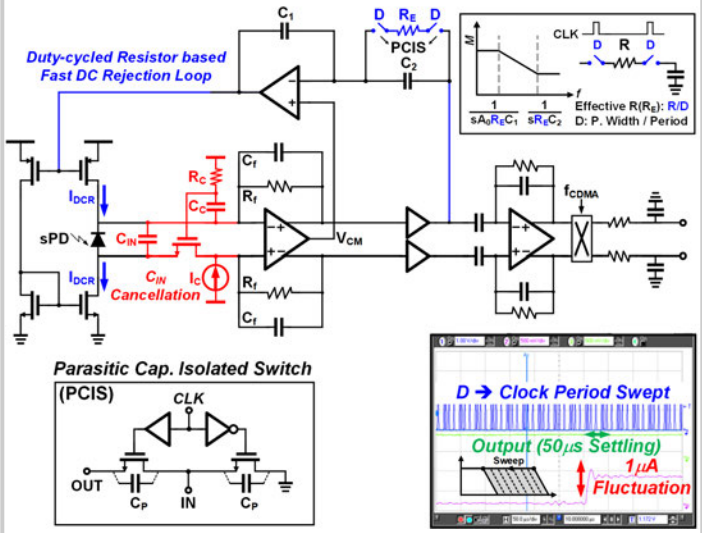


Figure 17.7.4: The circuit schematics and the measurement results of DOT receiver.

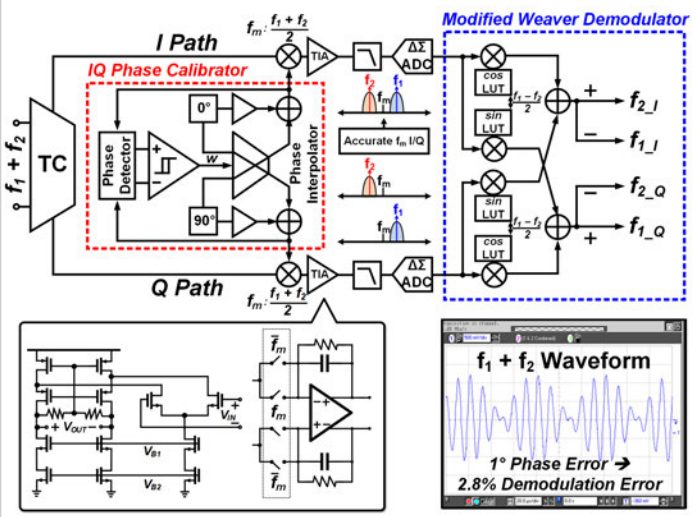


Figure 17.7.5: The architecture and schematics of EIT receiver.

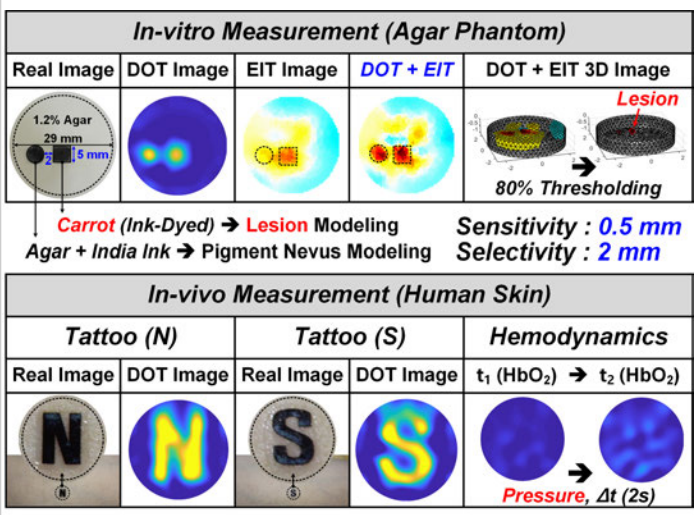
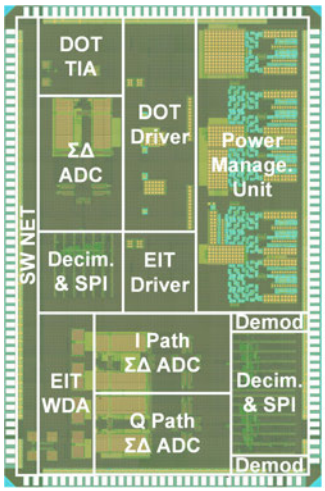


Figure 17.7.6: The measurement results of in-vitro and in-vivo environment.



Process	180 nm High Voltage CMOS	
Die Size	3.3 x 5.0 mm (Including Pads)	
Supply Voltage	1.8 V for V_{DD}, 7 V for V_{LED}	
Power Consump.	7.54 mW for Average	
Power Management Unit	Input Voltage	2.8 V ~ 4.2 V
	Peak Efficiency	Worst 51 %, Best 74 %
	Calibration Scheme	Coarse(5'b), Fine 2 Steps
CDMA DOT Current Driver	Current Range	0.05 ~ 6.4 mA
	Operation Freq.	Clk: 16 MHz, Chip Rate: 3.9 kHz
	Calibration Scheme	4'b Extinction Ratio Calib.
DOT Receiver	Frame Rate	7.0 fps
	Dynamic Range	100.4 dB (TIA DR)
	DCR Settling Time	50 µs
EIT Receiver / Current Driver	Frequency	f_C: 10 k ~ 80 kHz, f_I: 20 k ~ 100 kHz
	Calibration Scheme	I/Q Mismatch Calibrator
	Current Range	0.1 ~ 1 mA

	AFE 4403	TBCAS 08	TCAS 16	TBCAS 15	JSSC 16	This Work
Application	PPG, SpO_2	PPG	Optical Comm.	PPG	PPG	DOT
TIA Dynamic Range (dB)	105	29.8	45	57	89	100.4
DC Rejection (DCR) Type	Digital DCR	Analog DCR	Digital DCR	Digital DCR	Digital DCR	Fast Settled Analog DCR
DCR Automation	X	O	X	O	O	O
Settling Time (s)	N/A	1.68	N/A	100 m	N/A	50 µ
CMDR (dB)	20 dB (I_1~10_{pA})	54.5 dB (I_1~53.4_{pA})	N/A	N/A	N/A	> 50 dB

Figure 17.7.7: Chip micrograph, performance summary and TIA comparison table.

978-1-5386-8531-0/19/$31.00 ©2019 IEEE

17.8 A 2.6μW Monolithic CMOS Photoplethysmographic Sensor Operating with 2μW LED Power

Antonino Caizzone, Assim Boukhayma, Christian Enz

EPFL, Neuchâtel, Switzerland

Photoplethysmography (PPG) is a key technology allowing non-invasive monitoring of vital indicators such as heart rate (HR) and oxygen saturation (SpO_2). Today, the total PPG sensor power consumption is dominated by the few tens of mA of the LEDs driving current. Different solutions have been proposed to solve this bottle-neck, either by reducing the LED duty cycle [1-2] or by non-uniform sub-sampling [3]. A heart-beat-locked loop system that significantly reduces the LED power has been recently demonstrated [4]. However, this power reduction comes at the cost of more complexity since it requires a non-trivial heart beat prediction scheme and intrinsically hinders the full PPG wave representation.

In this work, we address the power issue by combining a high sensitivity photo-detector with an ultra-low noise and low power readout chain achieving the same signal-to-noise ratio (SNR) compared to conventional solutions, but at significantly lower LED power. We present a PPG sensor integrating an array of pinned-photodiodes (PPD), commonly used in CMOS imagers for achieving sub-electron noise [5]. The PPDs array is implemented on the same chip with the full analog front-end (AFE) including the analog-to-digital conversion (ADC). The full CMOS integration allows one to dramatically reduce the parasitic capacitance leading to a larger conversion gain and a lower noise. This approach also provides higher miniaturization and lower cost compared to traditional solutions with off-chip photodiodes. The use of an array additionally enables spatial averaging leading to further noise reduction. Consequently, the LED power can be reduced dramatically. A prototype has been implemented in a 0.18μm CMOS Image Sensor (CIS) process, achieving 4.6μW total power consumption, including 1.97μW LED power, at 1.38bpm HR average error.

The block diagram of the PPG sensor is shown in Fig. 17.8.1. It consists of a fully integrated chip embedding an array of PPDs, a passive averaging block, a switched-cap (SC) amplifier and an ADC. The photosensitive area consists of four clusters of 50 rows and 256 columns. The 50 pixels along the i-column are assembled in a macro-pixel (MP). The pixels of one MP share the same source-follower (SF) saving power while maintaining a reasonably low parasitic capacitance at the shared sense node (SN). The spatial averaging is done in two steps: first, a charge averaging on the shared SN within the same column and, second, a voltage averaging among the 1024 columns at the SFs output. This leads to a considerable shot and read noise reduction, together with a full pre-filtering of the PPG signal itself. Next, the PPG signal is amplified by a low noise programmable gain SC amplifier. In addition to the noise optimization, the chosen pixel achieves between 66% and 75% quantum efficiency (QE) for the selected LED wavelengths (Fig. 17.8.4) further reducing the LED power needed for achieving the target SNR [1].

Figure 17.8.2 shows the circuit implementation together with the timing diagram. The sense nodes, shared by 50 PPDs and corresponding to the SFs inputs, are first reset. In parallel, each individual PPD starts integrating the impinging light corresponding to the ambient light (AL). In order to precisely control the light-induced charge integration, the PPDs are first emptied by the sink switch TGs. At the end of the first integration phase, the generated photoelectrons are transferred, via the transfer gate TGt, to the SNs. The capacitance of each SN, shared by the 50 pixels along the same column, converts the integrated photoelectrons into a voltage and performs the charge averaging across the 50 rows. After the transfer is completed the SF output voltages are sampled on capacitors $C1_i$, via $S1_i$. Next, the SNs are reset again and the LED is pulsed on. As above, the PPDs precisely integrate the LED light superimposed to AL, and the related voltage level is sampled on $C2_i$, via $S2_i$. The power switch S_{SF} is only closed during this charge transfer and sampling phase for minimal power consumption. By closing $S3_i$ and $S4_i$ all the capacitors $C1_i$ and $C2_i$ related to the same sample are connected in parallel and share their charge, resulting into a voltage equal to the passive spatial average of the array pixel output samples. In addition, the large capacitor resulting from the parallel connection of multiple column-level capacitors acts as a large hold capacitor for the following stage. The full averaging operation comes with a shot noise variance reduction of 50×1024 and a read noise variance reduction of 1024. In order to extract the difference of the two averaged samples, corresponding to the AL and the AL plus LED light, a SC amplifier is operated as shown in Fig. 17.8.2: first, the averaged AL sample is stored in C3, via S5 and SAZ. Then, SAZ is opened and the charge stored in C3 is transferred to C4. At the closing of S6, the amplified difference between the two average values is obtained at the amplifier's output, leading to an efficient AL cancellation. The amplifier embeds a programmable gain, from 1 to 32 to adapt to different operating conditions. It then drives a 14b incremental ADC, through SF2. The power switches S_{Amp} and S_{ADC} are only closed at the beginning of the array averaging and opened after the digitization.

A die micrograph of the 180nm chip is shown in Fig. 17.8.7. The linearity of the light to digital conversion and the total noise measurements are shown in Fig. 17.8.3. The sensor is exposed to an LED shining at increasing driving current and the resulting output digital number is acquired. The programmable amplifier gain has been set to 8, which enables a wide range of emitting light conditions without saturation. With the exception of sub-mA LED operations, the chip shows ±3% non-linearity in the light to digital conversion. Note that this includes all the sources of non-linearity from the LED, the readout chain to the ADC. Figure 17.8.3 also shows that the output noise standard-deviation (STD) remains constant across the dynamic range, demonstrating the effectiveness of the shot noise reduction by spatial averaging. The total noise measured at the output of an off-chip 11 taps FIR low-pass filter is 3.1 DN_{RMS} in average corresponding to an input-referred noise as low as 0.68 e^-_{RMS} per PPD, thanks to the noise shaping introduced by the incremental ADC. Indeed, the noise measured directly at the output of the ADC corresponds to 9.43 DN_{RMS}. This accounts for all the noise components including the readout noise, quantization noise, shot noise and LED flicker noise.

An in-vivo acquisition of PPG has been performed in reflection mode on the index finger. Figure 17.8.4 shows the LEDs and sensor module used for the experiment, as well as samples of the PPG signal obtained using different LED wavelengths at a record-low LED average power of 1.97μW. Figure 17.8.5 shows a comparison between the HR directly extracted from the sensor output and a commercial ECG chest strap featuring a HR average error of 1.38bpm only measured on three healthy male subjects. These measurements have been obtained with a sampling frequency of 40Hz, an average LED driving power of 1.97μW, at a duty cycle of 0.07%, and a readout (AFE+ADC) average power consumption of 2.63μW only. This is obtained, referring to Figure 17.8.2, by closing S_{SF} for 80μs and S_{Amp} with S_{ADC} for 0.5ms.

Compared to recent state-of-the-art (SOA), shown in Fig. 17.8.6, this work achieves the lowest power consumption. Moreover, among the most relevant SOA works, this is the only solution featuring a full integration of the photosensitive area together with the AFE and the ADC. This last feature, in addition to the extremely low power consumption, makes this solution a serious candidate for the future of wearable digital healthcare.

References:

[1] K. N. Glaros, et al., "A Sub-mW Fully-Integrated Pulse Oximeter Front-End," *IEEE TBioCAS,* vol. 7, no. 3, pp. 363-375, June 2013.

[2] E. S. Winokur, et al., "A Low-Power, Dual-Wavelength Photoplethysmogram (PPG) SoC With Static and Time-Varying Interferer Removal," *IEEE TBioCAS,* vol. 9, no. 4, pp. 581-589, Aug. 2015.

[3] P. V. Rajesh, et al., "A 172μW Compressive Sampling Photoplethysmographic Readout with Embedded Direct Heart-Rate and Variability Extraction from Compressively Sampled Data," *ISSCC Dig. Tech. Papers,* pp. 386-387, Feb. 2016.

[4] D. H. Jang, et al., "A 43.4μW Photoplethysmogram-Based Heart-Rate Sensor Using Heart-Beat-Locked Loop," *ISSCC Dig. Tech. Papers,* pp. 474-476, Feb. 2018.

[5] A. Boukhayma, et al., "A Sub-0.5 Electron Read Noise VGA Image Sensor in a Standard CMOS Process," *IEEE JSSCC*, vol. 51, no. 9, pp. 2180-2191, Sept. 2016.

 978-1-5386-8531-0/19/$31.00

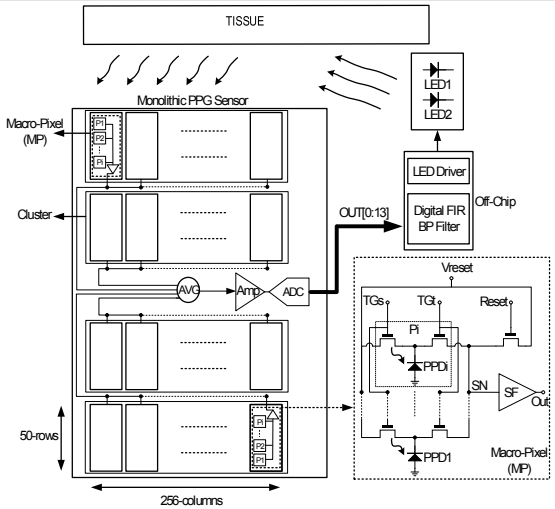


Figure 17.8.1: Block diagram of the monolithic PPG sensor embedding an array of macro-pixels, as photosensitive area, an averaging block, an amplifier and an ADC.

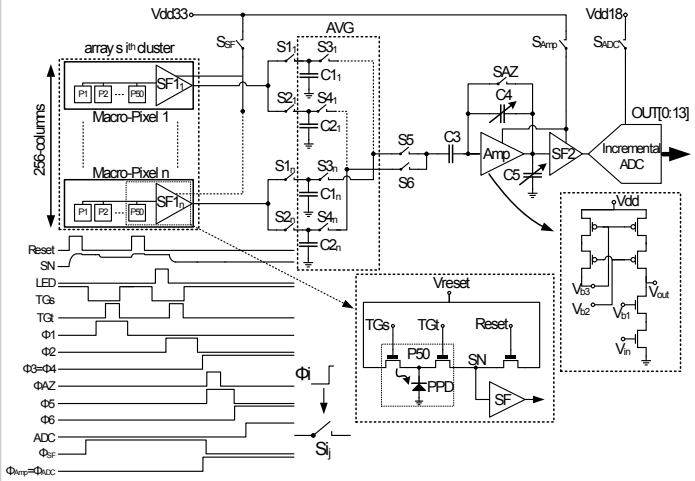


Figure 17.8.2: Architecture of the monolithic PPG sensor and detailed schematic of the OTA used in the amplifier, the averaging block, with the timing diagram.

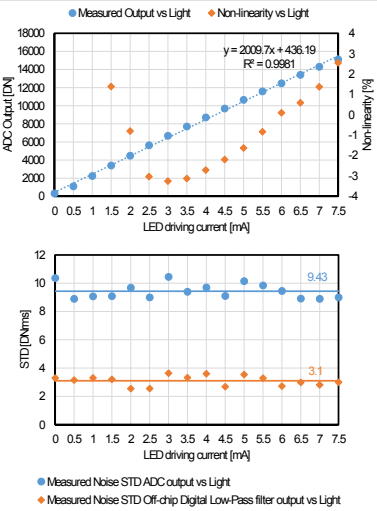


Figure 17.8.3: Measured light-to-digital linearity and total noise standard-deviation (STD), both at the ADC output and an off-chip FIR digital low-pass filter.

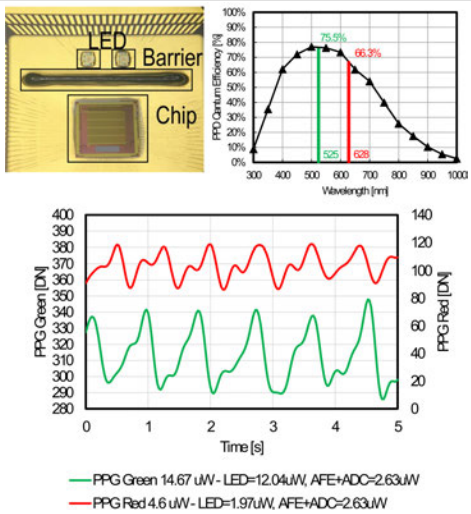


Figure 17.8.4: LEDs, chip and optical barrier. QE of the photodetector. PPG signals at two LEDs emitting power levels and colors and 2.63μW at the readout level.

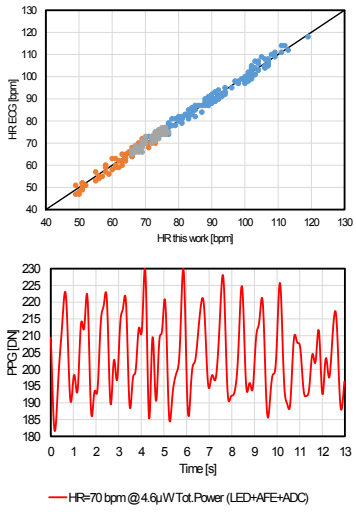


Figure 17.8.5: HR correlation with a commercial ECG chest strap. PPG signal of a subject. The experiment has been performed at 4.6μW total power (LED+AFE+ADC).

	This work	[2]	[3]	[4]
Year	2018	2015	2016	2018
Technology	180 nm	180 nm	180 nm	180 nm
Supply Voltage	3.3/1.8[1]	1.8	1.2	3.3
Sampling Frequency	40 Hz	165 Hz	4 Hz	100 Hz
LED Duty Cycle	0.07%	0.7%	0.0125%	0.0175%
Fully Integration[2]	Yes	No	No	No
Feature Extraction	HR/SpO2/RR[3]	HR	HR	HR
Avg HR error	1.38 bpm[4]	NA	NA	NA
Max HR error	3 bpm[4]	NA	10 bpm	2.1 bpm
Power LED	1.97 μW	120 μW[5]	43 μW[5]	16 μW[5]
Power AFE+ADC	2.63 μW	216 μW	172 μW	27.4 μW
Power TOT[6]	4.6 μW	336 μW[5]	215 μW[5]	43.4 μW[5]

1: ADC supplied at 1.8V.
2: Photosensitive area+AFE+ADC.
3: HR (heart rate), SpO_2 (oxygen saturation), RR (respiration rate).
4: Measured at 4.6 μW total power vs a commercial ECG chest strap on 222 separate measurements on three healthy individuals, corresponding to data from Fig.5.
5: Best case from the given data.
6: HR mode.

Figure 17.8.6: Comparison table.

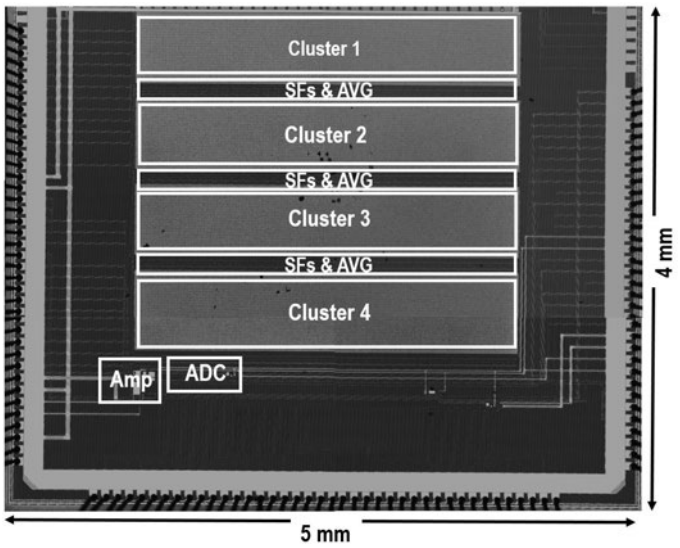


Figure 17.8.7: Micrograph of the fully integrated PPG sensor (4 mm × 5 mm).

978-1-5386-8531-0/19/$31.00